

# **For Reference**

---

**NOT TO BE TAKEN FROM THIS ROOM**



Ex LIBRIS  
UNIVERSITATIS  
ALBERTAEENSIS









T H E U N I V E R S I T Y O F A L B E R T A

RELEASE FORM

NAME OF AUTHOR ..... DOUGLAS G. BOUGHTON .....  
TITLE OF THESIS ..... DEVELOPMENT AND VALIDATION OF A .....  
..... CURRICULUM EVALUATION MODEL FOR .....  
..... THE VISUAL ARTS .....  
DEGREE FOR WHICH THESIS WAS PRESENTED ..... DOCTOR OF PHILOSOPHY .....  
YEAR THIS DEGREE GRANTED ..... 1976 .....

Permission is hereby granted to THE UNIVERSITY  
OF ALBERTA LIBRARY to reproduce single copies of  
this thesis and to lend or sell such copies for  
private, scholarly or scientific research purposes  
only.

The author reserves other publication rights,  
and neither the thesis nor extensive extracts from  
it may be printed or otherwise reproduced without  
the author's written permission.



THE UNIVERSITY OF ALBERTA

DEVELOPMENT AND VALIDATION OF A CURRICULUM  
EVALUATION MODEL FOR THE VISUAL ARTS

by



DOUGLAS G. BOUGHTON

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH  
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE  
OF DOCTOR OF PHILOSOPHY

IN

DEPARTMENT OF SECONDARY EDUCATION

EDMONTON, ALBERTA

FALL, 1976





Digitized by the Internet Archive  
in 2020 with funding from  
University of Alberta Libraries

<https://archive.org/details/Boughton1976>

THE UNIVERSITY OF ALBERTA  
FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled, "Development and Validation of a Curriculum Evaluation Model for the Visual Arts," submitted by Douglas G. Boughton in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Art Education.





## ABSTRACT

This study reports the development and validation of a curriculum evaluation model appropriate to the visual arts. Review of the literature indicated a lack of attention to evaluation issues in the field of art education. In addition evaluation models based upon "efficiency" concepts of education were viewed by art educators to be largely inappropriate for assessment of the visual arts.

The central proposition of the study was that an evaluation model derived from an artistic paradigm should be more suitable for the conduct of evaluative enquiry in the visual arts than existing models. The characteristics required of such a model were thought to be i) availability of an open ended frame of reference that does not demand prior specification of variables for selective analysis, ii) availability of interpretive techniques for the purpose of determining and describing educationally significant phenomena, iii) availability of methods of data collection and analysis that preserve the unique qualities of individual behaviours and events.

The concept of validity was examined with a view to exploring its meaning within alternate research paradigms. Empirical-analytical, hermeneutic, phenomenological, and artistic modes of investigation were reviewed. Two



purposes were served by this examination: i) to define the norms by which the model developed in this study would be validated and ii) to clarify the particular conceptions of validity appropriate for research methods employed by the model.

The model was developed by synthesising ideas derived from the literature and two curriculum development projects. A prototype of the model was tested on a conference type "mini" curriculum. Revisions following the pilot test produced the final form of the model. The analogy between art criticism and curriculum evaluation was explored at some length to explain the relationships of model components.

The validity of the model was demonstrated by attending to questions raised by the norms of correspondence, coherence, and pragmatism. Evidence from empirical and theoretical sources was employed to satisfy these norms. A description of an evaluation conducted for the Edmonton Public School Board in relation to an art appreciation project was used to illustrate the pragmatic value and the limitations of the model.

The model was found to possess a degree of validity given the limitations of the study. Further testing in a variety of situations was recommended.





## ACKNOWLEDGEMENTS

Production of this thesis would not have been possible without the assistance offered by many sympathetic individuals who performed contributory and supportive roles. I would particularly like to thank Dr. Ron MacGregor for his skilful and patient guidance and for the enormous amount of time and energy he expended on my behalf. I would also like to thank Dr. MacKay, Dr. McIntosh, and Dr. Harvey for their formative advice and encouragement.

The co-operation of Dr. Jim Simpson, Joan Dutton, Adrienne Coull, and Shirley Richardson is gratefully acknowledged. Without their support useful work in the field would have been impossible.

My admiration goes to Thelma who had to live with a husband who was frequently irascible and recalcitrant, and at the same time coped with the typing of all drafts of this document. Finally, thanks are due also to my mother for her invaluable contributions.





## TABLE OF CONTENTS

CHAPTER		PAGE
I	THE PROBLEM.....	1
	Introduction to the Problem.....	1
	Statement of the Problem.....	3
	Discussion of the Problem.....	4
	Purposes of the Study.....	14
	Definitions.....	16
	Delimitations.....	17
	Limitations.....	18
	Rationale for the Study.....	19
	References.....	21
II	REVIEW OF THE LITERATURE .....	23
	Evaluation Since 1900.....	23
	Summary of General Evaluation Theory (1900-1975).....	38
	Evaluation in Art Education.....	39
	Summary of Evaluation in Art Education....	51
	References.....	53
III	DESIGN OF THE STUDY.....	57
	Purposes of the Study.....	57
	Statement of Procedure.....	58
	Reference.....	65
IV	THE MEANING OF "VALIDITY".....	66
	Overview of the Concept.....	67



CHAPTER	PAGE
Theory Validation.....	68
Validity in the Empirical-Analytical Paradigm.....	71
Measurement.....	72
Experimental Design.....	78
The Meaning of Validity Within Alternative Research Paradigms.....	80
The Hermeneutic Tradition.....	86
The Phenomenological Tradition.....	91
The Sociological and Anthropological Traditions.....	96
The Artistic Tradition.....	97
Summary of Chapter IV.....	105
References.....	109
V THE MODEL.....	112
Introduction.....	112
PART I: DEVELOPMENT OF THE MODEL.....	112
Procedures Followed.....	112
Description of Data.....	113
The 1975 Fine Arts Council Conference....	114
The Edmonton Public School Board Art Appreciation Project.....	114
Salient Ideas from the Literature.....	115
Pilot Testing of the Model.....	120
Critical Analysis of the Pilot Study.....	123
Rationale Directing Model Revision.....	125
PART II: EXPLANATION OF THE REVISED MODEL...	128
Dimensions of the Revised Model.....	129



CHAPTER		PAGE
	Paradigmatic Criteria.....	129
	Curriculum Elements.....	130
	Qualitative Exposition.....	133
	Relationships of Model Components .....	134
	A Proposed Model for Art Criticism .....	135
	The Revised Curriculum Evaluation Model.....	138
	Methodological Assumptions.....	142
	Validity of Implied Research Methodologies.....	145
	Summary of Chapter V.....	148
	References.....	152
VI	VALIDATION OF THE MODEL.....	155
	The Norm of Correspondence.....	156
	The Norm of Coherence .....	168
	Pragmatic Norms.....	176
	Evaluation of the EPSB Art Appreciation Project.....	176
	Background to the Evaluation.....	176
	Purposes of the Evaluation .....	179
	Methodologies Employed for the Evaluation.....	184
	"Within Classroom" Methodology.....	186
	"Outside Classroom" Methodology.....	190
	Validation.....	190
	Design of the Evaluation.....	193
	Derivation of the Findings.....	196
	Presentation of the Findings.....	213





CHAPTER	PAGE
Generalizability of Findings.....	214
Pragmatic Value of the Model.....	217
Summary of Chapter VI.....	222
References .....	224
VII SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS....	226
Summary.....	226
Conclusions.....	229
Recommendations.....	241
BIBLIOGRAPHY.....	245
APPENDIX A. CURRICULUM EVALUATION MODEL FOR THE VISUAL ARTS - (PROTOTYPE).....	254
APPENDIX B. EXPLANATION OF EVALUATION QUESTIONS : THE 1975 FINE ARTS COUNCIL CONFERENCE....	256
APPENDIX C. CONCLUSIONS AND RECOMMENDATIONS: FINE ARTS COUNCIL CONFERENCE REPORT.....	261
APPENDIX D. BROAD OBJECTIVES OF "INTRODUCING WAYS OF SEEING".....	269
APPENDIX E. SELECTED SAMPLE OF LESSON OBJECTIVES AND ACTIVITIES FROM EPSB ART APPRECIATION PROJECT.....	271
APPENDIX F. SUMMARY OF CONCLUSIONS: EPSB "ART APPRECIATION" EVALUATION.....	277



## LIST OF FIGURES

FIGURE		PAGE
1	TIME FRAME SHOWING MAJOR FOCI IN DEVELOPMENT AND TESTING OF MODEL.....	59
2	MODEL FOR ART CRITICISM (TWO DIMENSIONAL WORKS).....	137
3	CURRICULUM EVALUATION MODEL FOR THE VISUAL ARTS.....	139



## LIST OF TABLES

TABLE		PAGE
1	CATEGORIZATION OF TOPICS DISCUSSED BY CURRICULUM COMMITTEES.....	158
2	JUDGE AGREEMENT EPSB CATEGORY I: AGENTS.....	160
3	JUDGE AGREEMENT EPSB CATEGORY II: CONTENT.....	160
4	JUDGE AGREEMENT EPSB CATEGORY III: IMPLEMENTATION.....	161
5	JUDGE AGREEMENT EPSB CATEGORY IV: OUTCOMES.....	161
6	JUDGE AGREEMENT EPSB CATEGORY V: RATIONALE.....	162
7	JUDGE AGREEMENT EPSB CATEGORY VI: RESOURCES.....	162
8	JUDGE AGREEMENT FAC CATEGORY I: AGENTS.....	163
9	JUDGE AGREEMENT FAC CATEGORY II: CONTENT.....	163
10	JUDGE AGREEMENT FAC CATEGORY III: IMPLEMENTATION.....	164
11	JUDGE AGREEMENT FAC CATEGORY IV: OUTCOMES.....	164
12	JUDGE AGREEMENT FAC CATEGORY V: RATIONALE.....	165
13	JUDGE AGREEMENT FAC CATEGORY VI: RESOURCES.....	165
14	JUDGE AGREEMENT: SUMMARY TABLE.....	166
15	RELATIONSHIP OF EVALUATION QUESTIONS AND MODEL ELEMENTS.....	185
16	RELATIONSHIP OF DATA TYPE AND SOURCES.....	195





TABLE

PAGE

17	CLASS A - VISITATIONS AND LESSONS OBSERVED.....	197
18	CLASS B - VISITATIONS AND LESSONS OBSERVED.....	198



## CHAPTER I

### The Problem

#### Introduction to the Problem

Evaluation within the field of art education has been almost completely ignored during the last two decades. Despite growing interest in programme assessment in other disciplines, only recently have there been some attempts by art educators to come to grips with the philosophic and pragmatic problems of evaluative activity in the arts.

Since its inception in 1959 Studies in Art Education has printed only fourteen articles that focussed upon curriculum evaluation, constituting only five percent of the total number of articles published in sixteen years. This statistic is even less impressive if one considers that, of the fourteen evaluation papers, five were short responses to an article by Pohland (1972) dealing with participant observation as a research methodology.

If Studies may be regarded as a reliable index of the concerns of art educators it can be hypothesised that interest in evaluation is minimal, or the issue has been deliberately avoided.

A survey of a dozen popular art education text-books, published since 1942, revealed a similar lack of



commitment to evaluation issues. Hubbard (1967) made a serious attempt to deal with concepts, function, problems, criteria and procedure for evaluation of art programmes. Eisner (1972) examined contexts for, and content of evaluation. As well, he proposed "art criticism" as a form of evaluation. Other authors treated the topic in a perfunctory manner or ignored it completely. Most observed the necessity for evaluation but contributed little to methodology or content.

Examination of papers published in Art Education during the last ten years provided some increasing evidence of advance toward a conception of evaluation consonant with the goals of art education. In 1974 one entire issue of this journal was devoted to evaluation, thus enabling (or stimulating) a few scholars in the field to air their views.

Wilson (1971) in the Handbook of Formative and Summative Evaluation of Student Learning made a major contribution to evaluation of art learning. He delineated content areas in art, identified important "art behaviours", and provided examples of testing procedures that could be used to measure these behaviours. His study is, however, a lonely one.

Generally, examination of the literature seems to indicate that most current models of evaluation are not congruent with the purposes of art education. The pronouncements of researchers in evaluation have, in the past, been largely ignored, or negatively received by art



educators.

### Statement of the Problem

The central question arising from the literature is related to a conception of curriculum evaluation that is appropriate for art education. Efficiency concepts of evaluation have created controversy in the arts because, as Steele and Victoria (1975) point out, a theory of accountability is lacking in art education. This omission has therefore committed the field to assessment by methodologies more suited to other disciplines (such as the sciences).

In light of this, the problem of this study is posed by the following question:

What is an appropriate model for curriculum  
evaluation in the visual arts?

Appropriateness in this context relates to congruence of methodological assumptions of the model and the epistemological framework of art education.

The definition of "model" employed for the purposes of this study is as follows:

An interpretive framework, for observation and analysis of a system of real events, that is characterized by a definition of dimensions and relationships operating among those dimensions.

Two elements are central to this definition - "dimensions" and "relationships" operating among these





dimensions. It is recognised that the dimensions identified within any curriculum evaluation model will be very much the same irrespective of the specific content area focus or its generic research style. However relationships among the dimensions will differ greatly according to the specific epistemological model.

### Discussion of the Problem

The methodologies of traditional evaluation models have historically leaned heavily upon measurement technology, particularly with respect to assessment of programme outcomes. But positivistic approaches to evaluative assessment in all curriculum areas have recently become the subject of increasing criticism. Alternative and complementary approaches have been suggested with increasing frequency by scholars, not necessarily art educators, who have become concerned about the anti-humanistic nature of hypothetico - deductive research methodologies.

In 1965 Bonner advocated a radical change in our view of the scientific method, away from statistical analysis toward a more introspective, self-anchored, phenomenological conception of research which he called "scientific humanism". His writing was influenced by Husserl's theories of phenomenology. Other research paradigms such as ethnomethodology, cognitive sociology, and hermeneutics have demonstrated growing significance for educational investigation in recent years (van Manen, 1974).



Three significant categories of inquiry processes were identified and discussed at length by Habermas in 1972. He called them "empirical - analytical", "historical - hermeneutic", and "critically oriented" sciences. Central interests of each of the categories are reflected in three epistemological traditions.

The "empirical - analytical" category employs a frame of reference that prejudges the meaning of possible statements and establishes rules, both for the construction of theories and for their critical testing. Habermas claims that empirical - analytical processes seek predictive knowledge. However, the meaning of such predictions (their technical exploitability) is established only by the rules according to which we apply theories to reality. Habermas goes on to suggest that the empirical sciences disclose reality subject to the constitutive interest in the possible securing and expansion, through information, of feedback monitored action.

The "historical - hermeneutic" sciences, on the other hand, gain knowledge from a different methodological framework. Access to the facts is provided by understanding of meaning rather than by observation. Knowledge in the hermeneutic tradition is always mediated through an interpreter's "pre-understanding". The understanding of meaning is directed in its very structure toward the attainment of possible consensus among actors in a framework of self understanding derived from tradition.

"Critically oriented" sciences of social action, such



as economics, sociology, or political science possess the dual goals of producing nomological knowledge and examining theory in action. The methodological framework is self-reflective, and is determined by an emancipatory interest.

This edited version of Habermas' explanation may serve to present to the reader a set of apparently arbitrary distinctions. His ideas, however, are provocative, particularly inasmuch as they provide a frame of reference for examination of present methodologies of education research. The first two categories, empirical - analytical and historical - hermeneutic provide the most useful comparison for the task at hand as the former is reflected in current evaluation models while the latter embodies some of the salient characteristics of "knowing" in art.

Within education during the last fifty years research methodologies borrowed from the natural sciences have proliferated. These methods have produced, within an empirical - analytical framework, a formidable body of knowledge to explain human behaviour. Discovery of nomothetic laws for prediction and control of human behaviour has been the primary purpose for this research. Apple (1974) observed that educational models of investigation have been drawn from behaviouristic sociology and psychology, fields purporting to pattern themselves after the strict sciences.

Evaluative research in education has drawn from the apparently successful methodological traditions of the main





body of educational research. Epistemological constraints that characterize researchers operating within the empirical-analytical paradigm have served to focus attention almost exclusively upon that which is quantitatively measurable in behavioural terms. Quality of the educational experience has been equated with efficiency in today's industrial, technocratic conception of education.

The recent push for accountability in public schools has served to re-inforce the industrial management style of assessing worth of educational programmes. Models displaying psychometric, economic and psychological biases have evolved to serve this need. Those with "systems" origins such as the Provus (1969) model are tied most closely to the industrial concept. Others, such as Stake's (1968) matrix are rather more flexible with respect to methodologies employed for data collection. Most have one element in common, that is, dependence upon explicit statements of objectives or needs, preferably in behavioural form. Methodologies borrowed from educational research are tailored to measure, among other things, levels of attainment with respect to prior stated objectives.

Educators' preferences for behavioural goals as a necessary pre-requisite for development and evaluation of curriculum reflects the underlying premise of the empirical-analytical category of research: that is, a frame of reference employed to pre-judge the meaning of possible statements derived from analysis of data. Moreover,



evaluation methodologies owe their constitutive interests in control and selective analysis of pre-determined variables (the stated behavioural goals) to their origins in the natural sciences.

To support the proposition that current concepts and methods of evaluation are incompatible with the purposes of art education it is necessary to juxtapose common values of art educators, with respect to research and purposes of art education, against the above analysis of evaluation epistemology. Art educators have historically displayed considerable reluctance to accept "scientific" research in their field.

Beittel (1960), Larson (1966), D'Amico (1966) and Eisner (1974) reported, in one form or another, reluctance of art educators to accept empirical - analytical methods of research or evaluation. Two major concerns were indicated by these writers. The first may be called a fear of "dehumanization" or loss of identity for the individual. The second is that the methods of art and science are antithetical, a point related to the problem of this study. From the perspective of an art educator these represent legitimate issues with respect to research, be it evaluative or elucidatory.

The fear of "dehumanization" is a by-product of the empirical - analytical tradition of research referred to before. Commonly used "scientific" methods of analysis in education rely greatly upon methods of group comparisons



rather than focussing upon individual behaviours. This concern reflects some important values underpinning the purposes of art education that can be traced to fundamental ways of knowing in art.

Embodied within Habermas' historical - hermeneutic category of research are the salient characteristics of "knowing" in art. The exclusive modality for knowledge acquisition in the hermeneutic tradition is mediation through an interpreter's "pre-understanding". Understanding of meaning is the central purpose. Mediation of artistic knowledge has, for several hundred years, been the institutionalized function of the critic. The critic, presumably, possesses skills of aesthetic criticism and articulation that served to make clearly visible, to those with more limited knowledge, the meaning embodied within artistic phenomena. Essential to the act is possession, by the critic, of prior extensive, and widely differentiated knowledge of artistic phenomena. The reader is referred to John Dewey (1934) for a most significant analysis of the process of aesthetic criticism.

The "critical" is not the only realm of activity in the arts. The act of producing an art object (although not unrelated to the process of criticism) is another commonly accepted method of inquiry for the purpose of gaining artistic knowledge. Such knowledge is generally idiosyncratic, although the artist may, at times, attempt to assume the role of critic in order to articulate the meaning of his work to a given audience. The most meaningful



knowledge gained from the productive act is tacit; i.e. derived from personal experience rather than by verbal communication from another.

The process of artistic production does, however, require critical skill in that a huge number of aesthetic decisions must usually be made throughout the act of creation. Each manipulative action requires subsequent artist consideration and judgement. Decisions range widely in importance according to the degree of change embodied in the final product. The primary mode of action by the artist is actively interpretive.

The epistemology of art is characterized by a concern to interpret qualitative meaning. Such interpretations are deliberately and explicitly subjective since they are the products of human interpreters. Habermas suggested that the understanding of meaning is directed in its very structure toward the attainment of possible consensus among its actors. However consensus is not a prerequisite for the understanding of meaning. It is manifestly impossible for any two interpreters to bring identical frames of reference to bear upon a given visual phenomenon.

Translating the concept "art" into an educational setting is an axiological problem that will never be satisfactorily resolved. Nevertheless the act of translation has invariably preserved the most significant elements of artistic knowledge. In effect, characteristics of the method of research in art have become the purposes of art education. Review of the literature indicates the following as





frequently suggested objectives for art education.

1. Understanding the meaning of visual phenomena through interpretive means, by way of productive or critical activity.
2. Understanding the relationship of art and culture.
3. Development of the visual aesthetic response.

To this point it has been suggested that current evaluation theory and the purposes of art education have developed within dissimilar research paradigms. The contention that most current methods of evaluation are inadequate for assessment of art programmes is supported, for various reasons, by a growing body of literature. The central conception of this argument is probably best reflected by Polanyi's (1959, 1964) suggestion that knowledge extends beyond mere objective (explicit) knowing to a substructure of tacit (subjective) knowledge. In light of Polanyi's work Apple (1974) observed,

that it may be necessary to reconsider the current emphasis in evaluation on explicit knowledge - an emphasis that can effectively destroy the act of personal knowing that Polanyi argues is the fundamental property of scientific and aesthetic awareness. In our stress on quantifiable achievement we may be negating the very element that makes anything worth knowing. ... If it is correct that explicit knowledge is less important than the process of tacit knowing, then our evaluation efforts are, to a significant extent, misdirected." (p. 25)

Evaluation models designed to serve the interests of efficiency in education are characterized by rigorous attention to quantitative measurement of human behaviours



that represent, by Polanyi's definition, objective knowledge. The purposes of art education, as was delineated above, are directed towards acquisition of tacit knowledge. Herein lies the nub of the problem, a dilemma that has been recognised by a number of art educators.

Efland (1974) observed that evaluation, directed towards assessment of degree of attainment of pre-specified behavioural objectives, was not in itself a worthless procedure. Rather, it is an activity that discovers information not central to the teaching of art. Efland also made the point that attitudes, values, and tastes take time to develop, very often coming to fruition years after teaching has ceased.

Eisner (1973, 1974) has compiled a lengthy list of limitations of the behavioural (efficiency) approach to curriculum development. These limitations, because of their nature, also directly affect conceptions of evaluation based on the behavioural approach. They are synthesised as follows:

1. Statements of educational goals describing mental events, feelings, attitudes, and values, not observable in manifest events, are ignored. The focus therefore becomes directed towards those aspects of human behaviour that are measurable.
2. Statements of specific objectives prior to the beginning of curriculum development have the effect of creating strategies of instruction that confuse



the logical with the psychological.

3. Behavioural objectives confuse the application of a standard with the making of a judgement in appraising outcomes.
4. Subject matters tend to be treated alike with respect to degree of specificity required.
5. Behavioural objectives do not take into account delayed and diffuse learning.
6. The degree to which it is possible to predict educational outcomes is overestimated by behaviouristic approaches.
7. Behavioural goals can only be stated in small units, therefore shifting the focus from larger and perhaps more important objectives.

Eisner (1975) acknowledged that although evaluative research based upon the assumptions and procedures of the natural sciences has a useful part to play in education it does not exhaust the ways in which men come to know. He suggested, as a needed complement to existing evaluative procedures, a "non - scientific" approach requiring educational connoisseurship and educational criticism.

Efland's and Eisner's observations are not isolated. Each of their concerns has been voiced entirely, or in part by others in the field. Stewart (1972) questioned group comparisons in evaluation and stressed the necessity for focus upon the individual. Hausman (1971) lamented situations in which items are counted just because they are



quantifiable, or complex and subtle distinctions are reduced to simple arithmetic categories. Kaelin (1969) criticised currently developing techniques of behavioural analysis because they fail to take into account the uniqueness of the individual. Bradley (1969) also criticised behavioural evaluation as arbitrary and uneducational.

The problem of this study has been identified in light of the controversy raised in the arts by efficiency concepts of evaluation. Lack of an adequate theory of evaluation in the arts has committed the field to assessment by inappropriate methodologies drawn from other disciplines.

### Purposes of the Study

The study has two major purposes:

1. To develop a curriculum evaluation model appropriate to the visual arts.

In this study it is proposed that the epistemological assumptions of an evaluation model for the visual arts should logically be drawn from an artistic (interpretive) paradigm rather than the traditional empirical - analytical mode of educational research.

Appropriateness, as was outlined in the Statement of the Problem, relates to congruence of the methodological assumptions of the model and ways of knowing through art.

More specifically, in light of points raised in the





Discussion of the Problem such a model should possess the following characteristics not demonstrated by empirical - analytical conceptions of evaluation:

- i) It should be independent of a frame of reference designed to pre-determine specific variables for selective analysis.
- ii) It should not depend primarily upon quantitative measurement and analysis technology for the purpose of establishing educationally significant phenomena. Rather the techniques of aesthetic criticism should be employed to this end ("connoisseurship" and "criticism" as defined on the following page).
- iii) It should be designed so that the unique qualities of individual or small groups of behaviours and events are not eliminated in favour of generalizable information.

2. To validate the model against theoretical and empirical references.

When viewed in relation to models (or theories) the concept of validity becomes quite complex. In this study the proposed model is measured against the norm of correspondence, the norm of coherence and pragmatic norms. (Kaplan, 1964)

Two distinct validation viewpoints are to be observed in this study.

- 1. Validation of the conception of the model in terms of the dimensions identified and the



relationships operating among those dimensions.

2. Validation of the evaluative information produced through direct application of the unique characteristics (listed above) of the proposed model.

Further elaboration and discussion of the concept of validity is provided in Chapters IV and VI.

### Definitions

Art Curriculum: All the experiences that learners have under the auspices and direction of the school, in the name of art education.

Aesthetic Criticism: An activity aimed at re-education of perception (Dewey, 1934). Description, analysis, interpretation, and evaluation are processes that may contribute towards this end following apperception, by the critic, of the phenomenon in its unified form.

Educational Connoisseur: One who is knowledgeable of the complexities of educational theory and practice and is, by interest and experience, capable of perceiving subtle characteristics and qualities within the educational milieu (Term coined by Eisner, 1975).

Educational Criticism: The process by which the educational connoisseur communicates his perceptions, using descriptive and expressive language, to his audience (Term coined by Eisner, 1975).

Empirical - Analytical Research: A method of



investigation that holds prediction and control as its central interests, employs a frame of reference that pre-judges the meaning of possible information and is directed towards discovery of nomothetic laws.

Evaluation: An assessment of worth.

Hermeneutic Research: A method of investigation directed towards understanding of meaning by a process of mediation through an interpreter's preunderstanding. (Habermas, 1972).

Model: An interpretive framework, for observation and analysis of a system of real events, that is characterized by a definition of dimensions and relationships operating among those dimensions.

Paradigm: A set of metatheoretical assumptions that determines for research, a methodology, focus, conceptual framework, type of data considered relevant and style of integration of new knowledge. (Adapted from Kuhn, 1962)

Qualitative Exposition: An appraisal of events or objects by a critic employing the principles of aesthetic criticism.

### Delimitations

The delimitations of this study relate to the nature and scope of the projects used to assist in developing and validating the model, teachers involved in the evaluation, and number of researchers available. These delimitations



are listed below.

1. The model was to be tested on one extended curriculum development project (The Edmonton Public School Board Art Appreciation Curriculum Development Project).

2. Pilot testing of earlier configurations of the model was conducted on a short conference type programme (The Fine Arts Council Conference, 1975).

3. Data were collected from all of the teachers involved in the "trial" stage of the EPSB project.

4. Schools in which the researcher tested the methods of educational connoisseurship and educational criticism were limited to two.

5. The author of this study was the only classroom observer.

### Limitations

The purpose of classroom observation for this study was to collect empirical data sufficient to allow consideration of the validity of the methodological assumptions of the model. Production of a full scale evaluation of the EPSB Art Appreciation Project was not viewed as the central purpose of this research. More extensive use of the methods of Educational Criticism and Educational Connoisseurship, (Eisner, 1975) would have served the ends of that evaluation (and this study) but was not possible because of:

1. Difficulty of entry. Curriculum materials were





being tested by a limited number of volunteer teachers on a trial basis and the project director was apprehensive of the effect of a widespread (external) evaluator intrusion.

2. Lack of funds. Classroom observations were conducted by the author of this study as no funds were available to hire additional qualified personnel. However, the author has had sufficient classroom experience to be fully cognizant of the variables attendant upon classroom observation.

Videotapes of the classrooms were viewed by another observer who is a graduate student in secondary art education with five years of teaching experience.

Again it is recognised that production of a full scale evaluation would require several observers (educational connoisseurs/critics) to generate the richest base of descriptive data. As well employment of several observers would permit visits to a larger number of classrooms. However, for the reasons stated in 1. above, one observer and one "videotape viewer" were considered sufficient for the purposes of the problem of this study.

### Rationale for the Study

Although a small nucleus of art educators has seriously worked towards the development of an acceptable conception of evaluation for the arts, the greatest proportion of the literature to date has been theoretical.



This study represents an attempt to synthesise some of the more recent notions of evaluative theory and translate them into a workable model for the visual arts.

Much of the scant evaluation literature in art education is devoted to the theme that empirical - analytical research methods, although effective for some purposes, do not uncover information central to art education. The "closed" framework of behavioural approaches is not consistent with the "open" concept of aesthetic learning.

Impetus for the study was generated by this idea and consequently the major contribution of this research lies in its exploration of the validity of an alternative approach.

Information presented herein may also serve to provide useful heuristic information for researchers wishing to explore further a concept of evaluation appropriate to the visual arts.



# REFERENCES

- Apple, M.W. The Process and Ideology of Valuing in Educational Settings. In M.W. Apple, M.J. Subkoviak & H.S. Lufner Jr., (Eds.), Educational Evaluation: Analysis and Responsibility. Berkely: McCutchan, 1974.
- Beittel, K.R. Art. In Chester W. Harris, (Ed.), Encyclopedia of Educational Research. New York: The Macmillan Company, 1960.
- Bonner, H. Two Views of Science. On Being Mindful of Man: Essay Toward a Proactive Psychology. Boston: Houghton Mifflin, 1965, 11-33.
- Bradley, W. Who Should Evaluate Art Students? Art Education, 1969, 22(9), 22-23.
- D'Amico, V. Millennium or Mirage. Art Education, 1966, 19(5), 27-32.
- Dewey, J. Art as Experience. New York: G.P. Putnam's Sons, 1934.
- Efland, A.D. Evaluating Goals for Art Education. Art Education, 1974, 27(2), 8-10.
- Eisner, E.W. Educating Artistic Vision. New York: The Macmillan Company, 1972.
- Eisner, E.W. Do Behavioural Objectives and Accountability Have a Place in Art Education? Art Education, 1973, 26(5), 2-5.
- Eisner, E.W. Toward a More Adequate Conception of Evaluation in the Arts. Art Education, 1974, 27(7), 3-5.
- Eisner, E.W. Some Alternatives to Quantitative Forms of Educational Evaluation. Thrust, 1975, 5(2). 13-15.
- Habermas, J. Knowledge and Human Interests. Boston: Beacon Press, 1971.
- Hausman, J.J. Curriculum Projects: Reflections on their Evaluation and Continuance. Art Education, 1971, 24(8), 24-27.
- Hubbard, G. Art in the High School. Belmont, California: Wadsworth Publishing Co. Inc., 1967.



- Kaelin, E.F. Are Behavioural Objectives Consistent with Social Goals of Aesthetic Education. Art Education, 1969, 22(8), 4-11.
- Kaplan, A. The Conduct of Inquiry: Methodology for Behavioural Science. Scranton, Pennsylvania: Chandler Publishing Company, 1964.
- Kuhn, T.S. The Structure of Scientific Revolution. Chicago: Chicago University Press, 1970.
- Larson, C.H. Search and Research. Art Education, 1966, 19(3).
- Pohland, P. Participant Observation as a Research Methodology. Studies in Art Education, 1972, 13(3).
- Polanyi, M. The Study of Man. Chicago: University of Chicago Press, 1959.
- Polanyi, M. Personal Knowledge. New York: Harper and Row, 1964.
- Provus, M. Evaluation of Ongoing Programs in the Public School Systems. The Sixty-eighth Yearbook of the National Society for the Study of Education. Chicago: The University of Chicago Press, 1969, Part II, 242-83.
- Stake, R.E. The Countenance of Educational Evaluation. Teachers College Record, 1968, 68, 523-540.
- Steele, C. and Victoria, J. Accountability and Assessment of Art: Education in Public Schools. Art Education, 1975. 28(5).
- Stewart, W. A Humanistic Study of the Visual Arts, The Commitment, Risk and Potential. Art Education, 1972, 24(2), 2-3.
- van Manen, M. Some Search Orientations for Research in Social Studies Education. Paper presented at the NCSS Research Advisory Committee Session, Chicago, 1974.
- Wilson, B.G. Evaluation of Learning in Art Education. In B.S. Bloom, J.T. Hastings & G.F. Madaus (Eds.), Handbook of Formative and Summative Evaluation of Student Learning. New York: McGraw Hill Publishing Company, 1972.





## CHAPTER II

### Review of the Literature

The following review of the literature is comprised of two parts. The first is a brief review of evaluation theory and significant models in the present century. In the second, examination of the literature pertaining directly to evaluation in art education is undertaken.

#### Evaluation Since 1900

Evaluation since the turn of the century can be divided into four periods demonstrating different emphases, and definitions of evaluation. The early 1900's to the 1930's was characterized by a conception of curriculum evaluation that equated administration of standardized tests with evaluation. Within this scheme the educator's primary concern was with content objectives.

In the early 1900's the work of Thorndike was instrumental in persuading scholars of the value of measuring change in human behaviour. The impetus of his scholarly activity was reflected in a burgeoning of measurement technology during this time.

Franklin Bobbitt (1924) in his book How to Build a Curriculum, posited a number of assumptions about curriculum development that marked the beginning of the



tradition of behaviourism later to be developed, in a curriculum context, by Tyler. Bobbitt suggested that all human life is made up of a number of specific activities that we all perform. Responsibility is upon the educator to somehow identify the nature of these activities and record them. The curriculum can then be constructed in such a way that experiences gained by the students will lead them along the way to facility in performing the previously identified "specific activities". These activities, therefore, are the goals of instruction. The measurement interests of Thorndike coupled with Bobbitt's suggestion about specification of objectives created fertile ground for growth of the "efficiency" movement in education of the 1910-1920's.

From the Thirties to the post Sputnik era, evaluation theory became somewhat more sophisticated. The basic emphasis here was upon formulation of objectives in curriculum development. Evaluation essentially consisted of assessing the "goodness of fit" between outcomes and objectives. Assessment of the nature of the objectives themselves was de-emphasised.

Ralph Tyler was the giant of the era. The Eight-Year Study (Smith & Tyler, 1941) was conducted along classic guidelines. Smith and Tyler formulated a philosophy for evaluation of this project that is still influential in the field at the present time. Important evaluation guidelines to come from the study were as follows:

1. Formulation of objectives. Emphasis was upon a



priori statements of objectives, preferably in behavioural terms.

2. Classification of objectives for the purpose of determining areas of emphasis and instruments to be used.
3. Definition of broad objectives.
4. Suggestion of situations in which achievement of objectives would be shown.
5. Selection of evaluation methods (measuring instruments).
6. Development and improvement of methods.
7. Interpretation of results, i.e. outcomes in relation to objectives.

The influence of Tyler's work is difficult to underestimate. Worthen and Sanders (1973) claim that Tyler has influenced evaluation for thirty years and continues to do so, as may be seen by the huge National Assessment Project that was begun in the United States in 1964. Moreover, notions embodied within the guidelines of the Eight-Year Study have been subsequently taken up and developed by other curriculum and evaluation theorists.

In the late fifties and early sixties monumental attempts to classify educational objectives within three domains - cognitive, affective, and psychomotor - were made by Bloom (1956) and Krathwohl, Bloom and Masia (1964). These attempts represented a move to refine Tyler's guidelines relating to specification, selection and classification of objectives.



Mager (1962) devised a scheme for explicit formulation of behavioural objectives. Any acceptable statement of objective, according to Mager, should: i) state the kind of behaviour desired; ii) specify the conditions under which it should occur; and iii) state the criteria for acceptable level of performance. Mager's work was instrumental in easing the task of evaluators interested in identification and measurement of desired outcome behaviours.

From the late fifties an important change was evidenced in evaluation thinking. Cronbach in 1963 urged that the traditional concept of evaluation be extended to course improvement and consideration of outcomes ranging far beyond the content of the curriculum itself. Cronbach also questioned the use of "comparative" evaluation, a technique that had been in common use from Thorndike's time. A switch in focus from outcome to process evaluation was subsequently evidenced in the early sixties following Cronbach's paper.

In the later sixties, a concerted effort was made by evaluators to formulate a universal conception of evaluation. Without question the most acclaimed work of the decade was Scriven's (1967) milestone article "The Methodology of Evaluation", that served to clarify several important issues. Scriven differentiated between "roles" and "goals" of evaluation. Moreover, he admonished evaluators for allowing roles to become the goals of their





activities. A second classical distinction he made was between "formative" and "summative" roles of evaluation. "Formative" evaluation is designed to feed back information to the curriculum developers for the purpose of improving the programme while it is still under development. "Summative" evaluation is usually aimed at the consumer and designed to assess the final product.

Scriven was instrumental in changing the flavour of evaluation in the sixties. The greatest impact of his work lay in his insistence that evaluative activity required the making of value judgements. Focus of evaluations subsequently shifted to include examination of the worth of objectives. In fact, Scriven, in a later paper (1972), extended this notion to extoll the virtues of "goal-free" evaluation. This concept implies that an evaluator is handicapped should he discover the developers' goals for any programme being examined. Scriven suggested that desirable goals for the programme should be determined by the evaluator, in the course of his investigations, by conducting a comprehensive needs assessment in the area.

Although the work of Cronbach, Scriven and others, served to clarify to a certain extent, the theory of evaluation in the sixties, certain pragmatic questions remained. A rather disparate collection of models evolved during that decade which represented attempts to operationalize current evaluation theories. Taba and Sawin (1962) and Walbesser (AAAS Commission on Science Education 1965) proposed models reflecting Cronbach's notion of



evaluation for course improvement. Significant among evaluation models proposed later in the sixties were those of Taylor and Maguire (1966), Stake (1967), Stufflebeam (1967), Provus (1969), Hammond (1969), and Alkin (1969).

Taba and Sawin criticised existing methods of evaluation. They noted five areas of deficiency that were widespread in curriculum evaluation designs. These deficiencies were:

1. The objectives which formed the basis for evaluations were usually too narrow.
2. The range of instruments and devices used to collect data was too limited.
3. The focus of evaluation was on the end product rather than the process.
4. Evidence from evaluation was often interpreted without adequate information about factors affecting learning and achievement.
5. There was no successful established method for translating evaluation findings into curriculum decisions.

To correct the above deficiencies Taba and Sawin prepared a model for evaluation that employed a wide range of instruments and techniques for data collection. Examples of significant evaluation information were suggested by the authors as methods of teaching, kinds of assignments given, patterns of pupil teacher interaction, equipment and facilities and administrative practices of the school.

The focus of this model demonstrated a switch



from outcome assessment to an examination of process. Heavy reliance was placed upon instrumentation to do this. Taba and Sawin reported upon the efforts of the Commission on Evaluation to compile a massive file of instruments and procedures useful for teachers to use routinely for the purpose of continuously monitoring the process and outcomes of instruction.

In a paper written for the Commission on Science Education of the American Association for the Advancement of Science, Walbesser (1965) described the design of the evaluation applied to an experimental elementary science curriculum. As the name, "Science - A Process Approach", implies, the focus of this curriculum was upon process. Consequently, the evaluation design, like the Taba and Sawin model, was adapted to examine intermediate components of the terminal programming outcomes.

To do this, behavioural elements of the Tylerian approach were refined and adapted. "Process" was regarded in this model as a hierarchy of behaviours ordered in a dependent sequence from the least to the most complex. Failure of a student to achieve a subordinate objective would explain his inability to acquire higher level behaviour.

The underlying premise of this evaluation procedure was that the curriculum project specify the class of reliably observable behaviour which could be expected to be acquired by individuals exposed to a particular set of instructional materials. The subsequent purposes of



the evaluation were i) to determine if these performances had, in fact, been acquired and ii) to assess the degree to which they were generalizable to new situations.

Neither the Taba/Sawin or Walbesser models questioned the first principles of the curriculum developers. Both methods were concerned only with examining process and outcomes rather than "educational value" of curriculum materials and procedures.

In 1963 Atkin recognised some of the shortcomings of the behavioural approach to evaluation that foreshadowed later criticism by Stake (1975), Eisner (1973a, 1974), Apple (1974) and others. Atkin pointed out that traditional evaluation procedures require curriculum developers to formulate clear statements of anticipated behaviours prior to implementation of the programme. The possibility that desirable and valuable behaviours, particularly those that were unanticipated, could be identified in retrospect was a procedure that was commonly regarded as too slipshod by most evaluators. Atkin recommended use of loosely structured classroom and student observation procedures for post hoc identification of programme objectives.

The Taylor and Maguire model was the first to include a value dimension. Measurement and assessment of value were the two basic evaluative activities posited by these authors. The measurement component examined outcomes, methods, goals, environment and personnel. Also, the relationships between these factors were examined. The value component comprised judgements of quality and





appropriateness of goals, strategies and outcomes.

The authors claimed that until a theory of evaluation is eventually formulated, practical evaluation will be facilitated by a model which sets up potential steps in the evaluation process. The Taylor-Maguire model is represented by four distinct judgmental stages i) broad objectives ii) interpretations iii) strategies and iv) outcomes. The authors qualify the representation of their model in this sequence with the observation that curriculum developers do not necessarily act in this manner. However, regardless of the actual schedule of events in developing a particular curriculum, the process can be thought of as having been developed in such a sequence.

The Provus, Alkin, and Stufflebeam models are closely related to management decision making models derived from systems theory. Provus' conception of evaluation was designed to serve as a continuous information management process that served programme implementation as well as programme assessment purposes. Like Tyler, Provus defined the goal of evaluation as a process designed to determine whether to improve, maintain, or terminate a given programme. In order to do this it was necessary for the evaluator to define programme statements, look for discrepancies between observations about the programme and pre-established standards, then feed back discrepancy information to the programme developers.



Although Provus (1969) identified three types of programmes with different patterns of development ("instant installation", "canned", and "school system designed" programmes) the major thrust of his work is directed towards evaluation of the first of these. The "instant installation" variety of programme can be described as one that has been quickly formulated without careful planning to utilize available resources.

To evaluate such a programme four major developmental stages were assumed. They were a) definition, b) installation, c) process and d) product. Provus described the process of evaluation as moving through stages and content categories in such a way as to facilitate a comparison of programme performance with standards, while at the same time identifying standards to be used for future comparisons.

The major shortcoming of the Provus model lies in its assumption of linearity in the curriculum development process. The prescriptive nature of its formulation serves to limit its flexibility in relation to its application in a variety of evaluation contexts.

The Stufflebeam and Alkin models displayed similar characteristics. Both employed decision management approaches and were strongly tied to curriculum development and implementation processes. Stufflebeam's model highlighted the various stages of programme development identified as context, input, process, and product.



The major objectives of the "Context" stage of Stufflebeam's CIPP model are to define the nature of the environment in which the change is to occur, discover the existing needs, and examine the problems underlying those needs. The "Input" stage of evaluation determines system capabilities, i.e. how to utilize resources to meet programme goals and objectives. During the "Process" stage continuous evaluation information is fed back to developers and administrators. The purpose of evaluation at this point is to detect or predict defects in procedural design or implementation. Stufflebeam observed that at this stage evaluation is multivariate and not all of the important variables can be specified beforehand. The evaluator should remain alert to unanticipated significant events. "Product" evaluation is used to determine the effectiveness of the project. Outcomes are related to objectives and Context to Input. Stufflebeam's description of method is as follows:

The method is to define operationally measured criteria associated with the objectives of the activity, to compare these measurements with predetermined standards on a comparative basis, and to make rational analyses of the outcomes using the recorded context, input and process information. (p. 131)

Alkin's model required almost identical procedures to Stufflebeam's, with the exception that separate foci must be employed for programme development and programme implementation. Five areas of evaluation were identified by Alkin. They were 1) Systems assessment, 2) Programme



Planning, 3) Programme Implementation 4) Programme Improvement and 5) Programme Certification.

Systems Assessment was viewed by Alkin as the means to determine educational objectives appropriate for particular situations. Programme Planning equates with Stufflebeam's "Input" stage in that information is provided to enable the decision makers to select among alternative processes. Programme Implementation evaluation determines the extent to which the implemented programme meets the description formulated by the decision makers during the planning stage. Evaluation for Programme Improvement is designed to provide as much information as possible about the relative success of aspects of the programme. The Programme Certification evaluation will enable global decisions to be made about the programme as a whole and its potential generalizability to other contexts.

Hammond (1969) fashioned the work of Bloom, Krathwohl, and Mager into an evaluation model for use by school district personnel. He suggested that the success or failure of innovations is determined by the interaction of specific forces found within the educational environment. These forces were conceptualized in a three dimensional space. An "Instruction" - "Institution" matrix was subdivided into organization, content, method, facilities, and cost under "Instruction" and student, teacher, administrator, educational specialist, family and community under "Institution". The third dimension was provided by a "Behaviour" dimension subdivided into the cognitive,





affective, and psychomotor domains. The interaction of variables from each of these three dimensions produced, in Hammond's view, all the possible factors to be considered in the evaluation of a given programme.

Hammond emphasised that use of the model in school districts required observation of five steps. The first is to isolate a specific subject area for examination rather than a number of disciplines (the total school programme). The second step is to define the descriptive variables in the instructional and institution dimensions. Stating objectives in specific behavioural terms (Mager, 1962) is the third step. Assessment of the degree to which behaviours, defined in step three, have been achieved is the prime requirement of step four. Finally an analysis of results within factors and relationships between factors will provide conclusions based upon behaviour.

The major distinction between Hammond's model and those of Provus, Alkin and Stufflebeam is that Hammond's is not linked closely to programme development and tends, as a result, to be a "static" tool rather than a dynamic means of viewing the evolutionary nature of curriculum in the educational context.

Stake (1968) devised an approach to evaluation that was somewhat less prescriptive, with respect to procedure, than the abovementioned plans. Two data matrices were employed for verification of descriptive and



judgemental data. Stake anticipated that an educator would gather data from a wide variety of sources in order to evaluate. These sources he roughly grouped into three potential data units - "antecedent", "transaction" and "outcome". He then divided the description matrix into two components, intents and observations while the judgement matrix was divided into standards and judgements.

Stake's model is the only one which clearly differentiates between subjective and objective procedures. The data net is cast widely in this approach, which implies that, although not much is missed in collection of evaluative data, depth and focus of the examination may suffer.

Stake has elaborated further upon methodology in a recent book called Evaluating the Arts in Education: a Responsive Approach (1975). The methodological framework he recommends in this volume clearly sets him apart from other evaluation theorists reviewed above. A "trade off" in measurement precision is advocated by Stake for the purpose of increasing usefulness of findings. He draws a distinction between pre-ordinate evaluators (i.e. those who place great significance upon goal statements, construct tests, predetermine standards, and rely upon comparisons of outcomes with objectives), and responsive evaluators. A responsive evaluator relies upon people relating naturally to events, is oriented more strongly to programme activities than programme intents, responds to



audience requirements for information, and reports different value perspectives of individuals or groups within the programme.

This later view of Stake's is reflective of the growing discontent with "efficiency" models of education and evaluation that characterizes much of the evaluation literature of the seventies. Apple (1974) suggested that evaluation is a process of social valuing, and, as such, requires choosing, from a range of competing value systems, one that gives meaning to educational processes or products. Consequently he sees evaluation as ideological in nature, that is, a taken for granted view of the educational situation which is shared by a number of people.

Karier (1974) identified problems of objectification, manipulation, choice, and freedom, raised by the "current efficiency craze in American education". (p. 300)

Instruments of the efficiency approach were identified by Karier as behavioural modification techniques, systems analysis, performance based instruction, and accountability.

Eisner (1973a) has accused educators of confusing efficiency with quality of education. He suggested that behavioural objectives are limiting tools for curriculum construction and evaluation. Two alternative types of objectives were offered to provide direction for educational activity without too closely defining the nature of the outcome. These objectives were called Expressive objectives and Type III objectives. An expressive objective merely



specifies the nature of the educational experience to be had, while the Type III objective delimits a specific educational problem to which there may be many acceptable, yet unpredictable, solutions. These kinds of objectives shift the emphasis in evaluation from careful specification of outcomes in advance, to a process of reflective analysis, conducted during and after the educational experience. In other words the assessment of "quality" of experience relies more upon judgement of the value of the experience than upon the efficiency of methods employed to gain predetermined goals.

#### Summary of General Evaluation Theory (1900 - 1975)

Between 1900 and the 1930's evaluation was equated with administration of standardized tests. Thorndike and Bobbitt were influential in generating interest in measurement of human behaviour, and statements of educational objectives in terms of specific behaviours, respectively. During the 1930's the Eight-Year Study, conducted by Tyler and associates, marked the beginning of a more rigorous period. Evaluation became, in simplistic terms, a comparison of outcomes with specifically stated behavioural objectives. Tyler's enormous influence in this tradition was felt for thirty years. The work of Bloom, Krathwohl, and Mager represented significant developments of Tylerian concepts. During and after the sixties, evaluation focus was shifted from outcome measurement and comparative evaluation, to process assessments. Evaluation became more strongly tied to





process of curriculum development. Scriven's influence was instrumental in persuading evaluators to make value decisions with respect to objectives of programmes previously accepted at face value. A variety of models also emerged during this period in response to the pragmatic problems of operationalizing evaluation theory. Most of these were Tyler influenced, measurement and systems oriented. In the early seventies discontent with efficiency concepts of evaluation has become increasingly apparent.

### Evaluation in Art Education

As was stated in the Introduction to the Problem, little has been contributed, by art educators, towards development of appropriate assessment technologies for the visual arts. Much criticism has been levelled at traditional methods of examining the arts in education. (D'Amico, 1966; Dobbs, 1972; Efland, 1974; Eisner, 1973a, 1973b, 1974; Hausman, 1963; Larson, 1966)

Contributions towards a theory of evaluation in the arts have been piecemeal. Little of consequence exists before the late fifties. Perhaps the most notable was the work of Smith and Tyler and the Evaluation Staff of the Eight-Year Study (1942). The Committee on Evaluation in the Arts listed three purposes of art teaching - appreciation, creative expression of experiences, and emotional adjustment. This work describes attempts to devise an instrument suitable for evaluation of appreciation. Claims made for



the instrument that was finally devised were i) that it could be used to ascertain changes in performance of students after they had taken art courses, and ii) that it could also be used for counselling students with respect to their ability to "appreciate" in relation to their peers.

Many of the early texts, and even some of the later ones, deal with "content" type evaluation of student's work while excluding programme assessment. "Content" evaluation is conducted by the teacher in individual, or group, situations, for the purpose of developing in the student sensitivities and abilities relating to the act of evaluation. This kind of evaluation is also justified on the grounds that it "reveals the child to himself", or "helps the child to clarify his thinking". (Conrad, 1964; Kundsens and Christensen, 1957; Lowenfeld, 1947; and more recently Dimondstein, 1974). "Content" evaluation is not strongly related to programme assessment, although apparent confusion with respect to this is obvious in the literature. Curriculum evaluation, as such, is largely ignored by the above authors.

De Francesco (1958) suggested six general aims of evaluation in art. They were:

1. To predict the degree of achievement in art.
2. To measure growth in concepts, material manipulation and expression.
3. To differentiate and certificate pupils.
4. To diagnose learning difficulties.
5. To diagnose defects in teaching procedures.



6. To determine remedial techniques.

No useful suggestions were offered by De Francesco regarding evaluation procedures to be used in order to achieve these aims.

This omission is a characteristic of most texts in art education. Although the necessity for evaluation is acknowledged, and at times strongly emphasised, methodology is almost always vague or absent.

Hubbard (1967) made a significant attempt to come to grips with some of the problems of curriculum evaluation in art education. He listed a number of "functions" of evaluation that echoed De Francesco. He then proceeded to outline problems of criteria specification, the difficulties of obtaining validity and reliability of judgements, and the shortcomings of teacher made tests. His orientation for assessment was essentially behaviouristic, with emphasis upon use of more than one judge for objectivity. His major focus was upon identification of "desirable" behaviours. The influence from Tyler was strong and obvious.

Wilson in 1968 proposed a "systematic curriculum evaluation scheme" for art education. His work reflected the strongest influence from the general evaluation theorists seen in art education to this time. Wilson's work demonstrated the effect of Stake, Tyler, and Scriven, in particular. His proposal can be reduced essentially to the five major checks he proposed as necessary for effective evaluation. They were:



1. Assessment of goal worth. The need for explicit criteria was expressed here.
2. Assessment of the relationships between goals and objectives. Objectives should be stated behaviourally. (No indication was provided relating to methodology for assessment of objectives not able to be stated in behavioural terms.) The logical contingency between goals and objectives should also be examined.
3. Assessment of the relationship of objectives to course content.
4. Relationship of course content to examination content, and the relationship of objectives to examination content. Wilson observed that there is a shortage of examination content in art for three reasons:

- i) The subject does not lend itself to conventional methods of measurement.
- ii) Art educators have not defined yet what they want to measure.
- iii) In some instances to attempt to measure is to change the nature of the phenomenon.

5. Assessment of objective achievement.

Wilson (1971), in a rather massive effort, developed his ideas more fully. His views changed slightly in the interim in that he took an apparently less rigid stand with respect to the necessity for statements of behavioural objectives. In fact he observed that it would be unfortunate





if all of the outcomes of a particular programme were completely specified in advance since the nature of art is open ended. On the other hand, he added, it would be unfortunate if no outcomes were specified prior to teaching, and no testing conducted. His work is significant, in this article, for, i) statements of content areas in art; ii) identification of art education behaviours; and iii) examples of testing procedures that were offered for each of the behaviours.

Three years later Wilson (1974) again expanded his view of evaluation of art instruction. His observations verged upon criticism of his own previous work. He suggested that the forces which shape art draw their impetus from the tensions operating between the contrary principles of order and chaos. However, the art that is taught in schools is largely a product of the principles of "order and convention". Wilson proposed that the behavioural, accountability and performance based movements in education, coupled with the conservatism that has historically dominated the field is likely to succeed in extracting any remaining pungency from the practice of art education. He then acknowledged his own contribution, in his previous work (1968, 1972a), to the "state of orderliness and the status quo".

Wilson claimed that this paper, entitled "The Other Side of Evaluation of Art Education", represented an attempt to draw attention to what was absent from the other articles,



i.e. the disorderly, illusive, aspects of art teaching and learning which defy controls and evaluation. To do this he identified and questioned the following three major groupings of assumptions that are apparent in the practice of art education.

1. That art is orderly and can be taught through a set of rules. In refuting this assumption Wilson drew upon Weitz' (1966) milestone article "The Nature of Art" in which the "open concept" of art was carefully developed. Because "art" is an open concept Wilson made the questionable claim that it is possible, to begin anywhere and go anywhere.

In summary he said

there is the danger that art teachers, their students, and society as well will conclude that schools should teach only orderly subjects capable of being broken apart, analyzed, and quantified. Actually what can be ordered, sequenced, behaviourally specified, measured and quantified is a pitifully small part of what ought to be known, felt, believed and produced in art. (p. 261)

2. That art education and art schooling are the same thing. Wilson claims a split exists between art education and art schooling. In fact, he proposed that for the art classroom to become effective in the future it will have to begin educating the way the culture does.

3. That art learning can be evaluated while students are still in school. Wilson challenged this assumption with the suggestion that the only important consequence of art education is the effective functioning of aesthetic



attitudes, beliefs, and values outside and beyond the school. This cannot be effectively assessed within the school context.

In conclusion Wilson made two final points. The first was that a strong need exists for sensitive searching out of unexpected outcomes of art schooling. These represent the most desirable consequences of instruction.

The second was that assessment methodologies appropriate for this task should be essentially open. In other words they should not be tied to any discipline, but would be most closely related to anthropology.

Lansing (1971) made a case for specification of "standards" in art education made up of specific goal statements. These standards should be established in relation to knowledge, attitude and skill. In fact, Lansing advocated establishment of national standards (in the United States) as most parents would prefer to know how far their youngsters had progressed in relation to other children in the country. Means of devising meaningful standards and useful methodologies of evaluation were not supplied. Lansing's only advice was that it must be done carefully.

Davis (1971) explained the role played by evaluation in construction of curriculum development packages under the auspices of CEMREL's Aesthetic Education Programme. The present emphasis, he suggested, was on formative evaluation to assist in programme development. The model for evaluation involved a broad definition of measurement, the



application of standards, and making of judgements. Certain behaviours were expected to be elicited from certain students by a particular package. Assessment devices were unrestricted and included oral questions, paper and pencil tests, and observational techniques. Summative evaluation was used to determine suitability of packages for various levels and results across teachers, schools and regions. Assessment techniques outlined by Davis reflect traces of Tylerism and acknowledge Scriven's formative-summative distinctions.

Development of the CEMREL model with particular emphasis upon the role of observational strategies was delineated in more detail by Smith (1972). Smith reported upon procedures used during the first year to "evaluate" CEMREL's Aesthetic Education pilot project. He described the valuable role played by descriptive narrative procedures for the purpose of developing concepts, hypotheses and problems to be studied in relation to the final evaluation model. The procedures used were adapted from Smith and Geoffrey's (1968) significant study Complexities of an Urban Classroom.

The data gathered in the first year contributed towards description and analysis of packaged instructional strategies and pupil aesthetic experiences. The ultimate aim of this analysis was to design detailed and carefully controlled performance outcome sub-studies.

A provocative article by Pohland (1972) on





"Participant Observation as a Research Methodology" sparked a number of responses. Pohland's article, although it did not report an evaluation conducted in an art context, appeared in Studies in Art Education with the implication that the methodology may be appropriate for the arts. Four dimensions, existing along a continuum, were suggested by Pohland to represent varying "styles" of participant observation. They were: i) emphasis on the descriptive narrative; ii) generation of theory; iii) verification of theory; and iv) quantification of data.

Pohland also identified two probable issues for participant observation as a methodology. The first was use by the evaluator of multiple observers, multiple instruments, and multiple variables in the research. Such a procedure was designed to provide a greater choice for construct validity. The second issue he called the "temporarily developing nature of naturalistic observation". By this he meant that one does not begin with a specific hypothesis and final design; instead, through observation, specific categories emerge gradually.

In response to Pohland, Stewart (1972) suggested that participant observation would be useful, both in formative and summative evaluation, despite the "observer bias" problem. Lewis (1972) countered Pohland's claims for the method by suggesting that flexibility and reliability were incompatible goals. Wilson (1972b) expressed interest in the notion that anthropological and sociological



methodologies be used (as in participant observation) for evaluation. This approach, he felt, was particularly appropriate for research in art education. His reasons were: i) the researcher is able to examine a greater number of variables and their interrelations concurrently; ii) the researcher can switch or devise new methodology in midstream if necessary; iii) the qualitative relationship required to exist between the researcher and the situation he is examining is appropriate for art education.

Elements of the methodology of participant observation were recommended by Dobbs (1972) who criticised emphasis upon outcome performance in evaluation activity in the arts. Instead, he suggested a shift in focus to process oriented assessment requiring the researcher to participate more fully in the programme and to draw from a spectrum of disciplines for research methodologies.

Michael Day (1972) observed that one of the most valued aspects of art education was preservation of the independence, individuality, and diversity of teaching style among art teachers. One of the factors that contributes to the diversity of art programmes is the relative lack of specific curriculum constraints and responsibilities. Lack of specificity is deliberate in art for the purpose of preserving its open ended nature. Consequently, Day claimed, the individual art teacher's rationale plays a critical role in determining the purposes,



direction and emphasis of the art programme. Evaluation of art programmes could therefore be greatly assisted by examining rationales of art teachers. Day developed an instrument to this end called the Day Art Rationale Assessment Instrument.

The contributions of Eisner to evaluation have been reported elsewhere in this review of the literature. Although Eisner is an art educator the major thrust of his work is more appropriately examined in relation to the field of general evaluation theory.

A powerful idea explored by Eisner (1973b) in relation to research (and consequently evaluation) in art education has to do with the difficulty of determining satisfactory criteria for performance in the visual arts. He suggested that there are a variety of skills which may be defined and assessed with precision but they are peripheral rather than desired outcomes in the field. He observed further that,

Determining the artistic outcomes of instruction requires the making of a judgement rather than the application of a standard, for there are no necessary and sufficient conditions for producing a work of art, appreciating it or understanding its place in history. Thus the problem of appraising effective art teaching, insofar as it is based upon the character and quality of student artwork, student response to art and student comprehension of its significance, is fraught with much difficulty, much more difficulty, in my opinion, than one finds in other fields. (p. 1199)

A concomitant component of this problem was identified by Eisner as the ineffability of the qualities



that constitute excellence in art. To discover and describe these qualities in the context of art education using psychological research paradigms is a particularly difficult undertaking. Eisner, in 1973, advocated development of new models for the study of qualitative aspects of teaching.

Two years later Eisner (1975a, 1975b) presented a proposal for a "complementary" model of evaluation that employs the methods of "educational connoisseurship" and "educational criticism". The primary assumption of this approach is that the purposes of evaluation are i) to improve instruction and ii) to provide information to the public about the quality and character of programmes. This can best be done by employing connoisseurship (the art of appreciation) and criticism (the art of disclosure) in the classroom. Eisner claimed that this approach enables discovery of the character of teaching, significance of the content, the nature of the students, the tasks engaged in, the quality of ideas expressed, the character of the products produced, the type of dialogue engaged in and the type of classroom atmosphere. Also he suggested the "richer" approach to presentation of evaluative information was more likely to promote change for the better in instruction. Although Eisner suggested this model in the broad context of general evaluation theory its epistemology is drawn from the artistic paradigm. Consequently it is particularly relevant.





Stake's (1975) conception of the "responsive evaluation", reported in the first section of this chapter, is reflective of Eisner's (1975a) vision of a more "artistic, catholic, eclectic" approach to evaluation. In his paper "To Evaluate an Arts Programme" Stake presented a methodology for evaluation that is essentially open ended. No specific variables or methods of observation were proposed prior to commencement of evaluative activity. Instead Stake posited a "clock" model of prominent recurring events that will take place during an evaluation. These events include such items as "conceptualize issues and problems"; "select observers, judges, instruments (if any)"; "thematize, prepare portrayals, use studies", etc. The strength of this method lies in its flexibility. Emerging issues are identified and dealt with throughout the evaluation. The emphasis upon "understanding" and discovery of programme "values" reveals a similar philosophy of evaluation to that espoused by Apple (1974).

#### Summary of Evaluation in Art Education

Contributions to evaluation by art educators have been piecemeal. Considerable criticism of traditional methods of evaluation has been recorded. Influence of writers in general evaluation may be seen in the work of Hubbard, Wilson, Lansing, and Davis. Primary sources of this influence were Tyler, Stake and Scriven. Some confusion exists in art education texts with respect to evaluation



as "content" and evaluation designed to assess programme worth. As well suggestions are limited regarding evaluation methodologies. Pohland proposed participant observation as an evaluation research methodology. Dobbs urged for a change in focus from outcome to process oriented evaluation in art education. Eisner suggested a model of evaluation drawn from the artistic paradigm that requires use of educational "connoisseurship" and "criticism". Stake has developed a more expansive open ended conception of evaluation for arts programmes called "a responsive approach".



# REFERENCES

- AAAS Commission on Science Education. An Evaluation Model and its Application. Washington D.C.: American Association for the Advancement of Science. Publication 65-9, 1965.
- Alkin, M.C. Evaluation Theory Development. Evaluation Comment, 1969, 2(1), 2-7.
- Atkin, J.M. Some Evaluation Problems in a Course Content Improvement Project. Journal of Research in Science Teaching, 1963, 1, 129-132.
- Apple, M.W. The Process and Ideology of Valuing in Educational Settings. In M.W. Apple, M.J. Subkoviak, & H.S. Lufler Jr., (Eds.), Educational Evaluation: Analysis and Responsibility. Berkely: McCutchan, 1974.
- Bloom, B.S. Taxonomy of Educational Objectives: Cognitive Domain. New York: David McKay Co., Inc., 1956.
- Bobbitt, F. How To Make a Curriculum. Boston: Houghton Mifflin, 1924.
- Conrad, G. The Process of Art Education in the Elementary School. Inglewood Cliffs, N.J.: Prentice Hall Inc., 1964.
- Cronbach, L.J. Evaluation for Course Improvement. Teachers College Record, 1963, 64, 676-683.
- D'Amico, V. Millennium or Mirage. Art Education, 1966, 19(5), 27-32.
- Davis, D.J. Human Behaviour: Its Implications Curriculum Development in Art. Studies in Art Education, 1971, 12(3), 10-12.
- Day, M. Rationales for Art Education: Thinking Through and Telling Why. Art Education, 1972, 25(2), 17-20.
- De Francesco, I.L. Art Education, Its Means and Ends. New York: Harper Brothers Publishers, 1958.
- Dimondstein, G. Exploring the Arts with Children. New York: McMillan Publishing Co. Inc., 1974.
- Dobbs, S.M. Program Analysis in Arts Education. Studies in Art Education, 1972, 12(2).



- Efland, A.D. Evaluating Goals for Art Education. Art Education, 1974, 27(2), 8-10.
- Eisner, E.W. Do Behavioural Objectives and Accountability Have a Place in Art Education? Art Education, 1973, 26(5), 2-5
- Eisner, E.W. Research on Teaching the Visual Arts. In M.W. Travers (Ed.), Second Handbook of Research on Teaching. Chicago: Rand McNally College Publishing Co., 1973b, 1196-1209.
- Eisner, E.W. Toward a More Adequate Conception of Evaluation in the Arts. Art Education, 1974, 27(7), 3-5.
- Eisner, E.W. Some Alternatives to Quantitative Forms of Educational Evaluation. Thrust, 1975a, 5(2), 13-15.
- Eisner, E.W. The Perceptive Eye: Toward the Reformation of Educational Evaluation. (Unpublished mimeographed paper of an address presented at a seminar, University of Alberta)
- Hammond, R. Context Evaluation of Instruction in Local School Districts. Educational Technology, 1969, 9(1), 13-18.
- Hausman, J.J. Research on Teaching the Visual Arts. In N.G. Gage (Ed.), Handbook of Research on Teaching. Chicago: Rand McNally, 1963, 1101-1117.
- Hubbard, G. Art in the High School. Belmont, California: Wadsworth Publishing Co. Inc., 1967.
- Karier, C.J. Ideology and Evaluation: In Quest of Meritocracy. In M.W. Apple, M.J. Subkovic, & H.S. Lufner Jr., (Eds.), Educational Evaluation: Analysis and Responsibility. Berkely: McCutchan, 1974, 279-319.
- Knudsen, E.H., Christensen, E.M. Children's Art Education. Peoria, Illinois: Chas. A. Bennett Co., Inc., 1957.
- Krathwohl, D.R., Bloom, B.S., & Masia, B.B. Taxonomy of Educational Objectives: Affective Domain. New York: David McKay Co., Inc., 1964.
- Lansing, K.M. Art, Artists and Art Education. New York: McGraw-Hill, 1971.
- Larson, C.H. Search and Research. Art Education, 1966, 19(3).
- Lewis, H.P. Point. Studies in Art Education, 1972, 13(3).





- Lowenfeld, V. Creative and Mental Growth. New York: The Macmillan Company, 1947.
- Mager, R.F. Preparing Instructional Objectives. Palo Alto, California: Fearon Publishers, 1962.
- Pohland, P. Participant Observation as a Research Methodology. Studies in Art Education, 1972, 13(3)
- Provus, M. Evaluation of Ongoing Programs in the Public School Systems. In R.W. Tyler (Ed.), The Sixty-eighth Yearbook of the National Society for the Study of Education, Part II, Chicago: The University of Chicago Press, 1969, 242-83.
- Scriven, M. The Methodology of Evaluation. In R.E. Stake (Ed.), Curriculum Evaluation. American Educational Research Association Monograph Series on Evaluation, 1, Chicago: Rand McNally, 1967.
- Scriven, M. Prose and Cons about Goal-Free Evaluation. Journal of Educational Evaluation, 1972, 3(4), 1-4.
- Smith, E.R., & Tyler, R.W. Appraising and Recording Student Progress. New York: Harper and Row, 1942.
- Smith, E.R., Tyler, R.W., & the Evaluation Staff. The Evaluation and Appreciation of Art. Adventure in American Education 3, New York: Harper & Brothers, 1942.
- Smith, L.M. The use of Observational Strategies for the Development of a Summative Evaluation Model. Washington University and CEMREL, Inc. Paper presented as part of the Symposium Evaluation Efforts in Laboratory and R & D Centres: Some Emerging Issues. AERA, Chicago, April 7, 1972. (ERIC Document Reproduction Service No. ED064367)
- Smith, L.M., & Geoffrey, W. The Complexities of an Urban Classroom. New York: Holt, Rinehart, & Winston, 1968.
- Stake, R.E. The Countenance of Educational Evaluation. Teachers College Record, 1968, 68, 523-540.
- Stake, R.E. To Evaluate an Arts Program. In R.E. Stake (Ed.) Evaluating the Arts in Education: A Responsive Approach. Columbus, Ohio: Charles E. Merrill Publishing Company, 1975. 13-32.
- Stewart, W. Point. Studies in Art Education. 1972, 13(3).



- Stufflebeam, D.L. The Use and Abuse of Evaluation in Title III. Theory into Practice, 1967, 6(3), 126-33.
- Taba, H., & Sawin, E.I. A Proposed Model in Evaluation. Education Leadership, 1962, 20, 57-71.
- Taylor, P.A., & Maguire, T.A. A Theoretical Evaluation Model. Manitoba Journal of Educational Research, 1966, 1, 12-17.
- Weitz, M. The Nature of Art. In D.W. Ecker & E.W. Eisner (Eds.), Readings in Art Education, Waltham, Mass.: Blaisdell, 1966, 49-56.
- Wilson, B.G. A Proposal for Systematic Curriculum Evaluation in Art Education. Studies in Art Education, 1968, 9(3).
- Wilson, B.G. Evaluation of Learning in Art Education. In B.S. Bloom, J.T. Hastings, & G.F. Madaus (Eds.), Handbook of Formative and Summative Evaluation of Student Learning. New York: McGraw Hill Publishing Company, 1972a, 499-558.
- Wilson, B.G. Point. Studies in Art Education, 1972b, 13(3).
- Wilson, B.G. The Other Side of Evaluation of Art Education. In G.W. Hardiman & T. Zernich (Eds.), Curricular Considerations for Visual Arts Education: Rationale Development and Evaluation. Champaign, Illinois: Stipes Publishing Company, 1974, 247-76.
- Worthen, B.R., & Sanders, J.R. Educational Evaluation: Theory and Practice. Worthington, Ohio: Charles A. Jones Publishing Company, 1973.



## CHAPTER III

### Design of the Study

#### Purposes of the Study

The problem with which this study is concerned is derived from the following question:

What is an appropriate model for curriculum evaluation in the visual arts?

"Appropriateness" was defined elsewhere in this study as congruence of methodological assumptions of the model and ways of knowing through art.

The above question was translated into two purposes for the study.

1. To Develop a curriculum evaluation model appropriate to the visual arts.

The characteristics possessed by such an evaluation model were thought to be i) Availability of an open ended frame of reference that does not demand prior specification of variables for selective analysis.

ii) Availability of interpretive techniques (such as connoisseurship and criticism) for the purpose of determining and describing educationally significant phenomena.

iii) Availability of methods of data collection and analysis that preserve the unique qualities of



individual behaviours and events.

2. To validate the model against theoretical and empirical referents.

Validity was felt to be a complex concept which should be considered, in this study, from two major viewpoints.

i) For the purpose of validating the model itself. In this case validation involved successful application of the norms of correspondence, coherence, and pragmatism. (Kaplan, 1964).

ii) For the purpose of explaining the concepts of validity employed by the methodologies implicit within the model. The overall acceptability of the model would presumably be enhanced by showing that the procedures employed for data collection and analysis were themselves valid.

Statement of Procedure

Although the purposes of the study have been stated separately, the actual processes of development and validation very often involve simultaneous procedures. For example, the nature of the procedures employed to define the dimensions of the model also determined the validity of those dimensions.

In Figure 1 a time frame showing the major foci related to development and testing of the model is presented. This figure illustrates concurrent and sequential interests









as these became important in the development of the model. Also, the length of time occupied by each focus is presented.

Ten major foci were identified. They were "Literature Review", "Development of Prototype of Model", "Revision of Model", "Observation of FAC Meetings", "Generation of FAC Evaluation Questions", "Observation and Analysis of FAC Conference", "Observation of EPSB Meetings", "Generation of EPSB Evaluation Questions", "Content Analysis of EPSB Materials", "Classroom Observation and Analysis EPSB". It should be recognized that there was no direct correlation between length of time devoted to each focus and the relative importance assumed by each focus within the total time frame.

Various conceptions of validity were considered to be appropriate to different aspects of the model. "Curriculum evaluation model" is a concatenation of intricate concepts that reflects a philosophic rationale for specific kinds of widely differentiated procedures. The model itself is representative of theories of "curriculum" and "evaluation" connected in one particular relationship. Validity of this relationship was anticipated to be appropriately demonstrated by applying three general norms for theory validation i) the norm of correspondence, ii) the norm of coherence and iii) pragmatic norms.

The model also takes on strength from recognition of the validity of the "widely differentiated procedures"



which the model isolates. These procedures may be described as the methodologies for data gathering and analysis; methodologies which may be anticipated to vary according to the nature of the data sought. For example, methods employed to gather and examine data relating to the products of art instruction would be subsumed by concepts of validity different from those governing the collection and examination of information about the facilities necessary for effective implementation of the curriculum. Explication of appropriate conceptions of validity (of the observations required by the model) was therefore considered to be a necessary part of the study.

Following is a chapter by chapter description of the procedures employed to realize the purposes of this study.

Chapter IV. This chapter consists of an exploration of the meaning of the concept "validity" within competing research paradigms. Two major benefits are to be derived from this exploration:

1. The parameters for validation of the model developed in this study (i.e. the norms of correspondence, coherence and pragmatism) are discussed.
2. The meaning of validity within alternative research paradigms is clarified so that a reference is provided for explanation of the research methodologies implied by



the model. This then makes possible clear definitions of the validation procedures employed in the "testing situation" described in Chapter VI.

Chapter V. The procedures employed to develop the model are described in the first part of this chapter. The revised model is presented in the second section.

Part I    Development of the Model

- i) Procedures followed.
- ii) Description of data sources and salient ideas which gave rise to the form of the model.
- iii) Pilot testing of the original model.
- iv) Rationale for development of the revised model.

Part II    Explanation of the revised Model

- i) The dimensions of the model.
- ii) The relationships operating between dimensions.
- iii) Methodological assumptions of the model.
- iv) Conceptions of validity appropriate to the methodologies employed by the model.

Chapter VI. Validation of the model is described in this chapter. Empirical and theoretical dimensions are covered within the validation procedure. Central questions





raised by the three general norms for theory validation are addressed individually in this chapter.

### 1. The Norm of Correspondence.

Does the theory fit the facts? As was mentioned before, the concept of "curriculum evaluation model" is a complex combination of ideas. To address satisfactorily the question stated above, these ideas must be separated. The marriage of three essential notions was viewed to constitute the structure of the concept. They were, theories of curriculum, curriculum development, and evaluation. Of these three, only curriculum development is demonstrable in an empirical fashion. "Curriculum" and "evaluation" are problematic abstract concepts, whereas "curriculum development" is a process which may be observed empirically. Consequently, the only aspect of the model which could be demonstrated to "fit facts" as such was the curriculum development dimension. Evidence was therefore sought to establish this congruence.

### 2. The Norm of Coherence.

- i) How does the model integrate with existing theories of curriculum evaluation?
- ii) In what respects does the model differ from existing models?
- iii) Is the logic of the model internally consistent?

Examination of the literature was expected to provide material for answers to these questions.



### 3. Pragmatic Norms.

- i) How well does the model function?
- ii) Does it serve evaluative purposes?
- iii) What is the nature and usefulness of the information generated by the unique characteristics of the model?

Answers to these questions were provided by examples of findings produced in the course of conducting an evaluation for the Edmonton Public School Board. It should be noted that the purpose of conducting this evaluation was to determine the usefulness of the unique characteristics of the model (such as use of an interpretive framework, and the procedures of connoisseurship and criticism in classroom observation). Results produced by traditional, time - tested methodologies were not emphasised.

Chapter VII. Conclusions arising from the material presented in the previous chapter are drawn. Implications for use of the model are also discussed.



REFERENCE

Kaplan, A. The Conduct of Inquiry: Methodology for Behavioral Science. Scranton, Pennsylvania: Chandler Publishing Company, 1964.



## CHAPTER IV

### The Meaning of "Validity"

The root meaning of "validity" is derived from a Latin term meaning strength. Its meaning in the context of research is related to knowledge. Strength of knowledge is usefully interpreted to mean that which provides the most effective means to the given ends. (Adapted from Kaplan 1964, p. 198) This rather nebulous interpretation connotes highly differentiated meanings according to the ends of the research in question.

The purpose of this chapter is to provide an overview of the many possible conceptions of the meaning of validity in competing and analogous research paradigms. This will then provide a backdrop against which the evaluation model developed in this study can be examined for the purpose of clarifying its claim to validity (or, as shall be argued later, "validities").

The concept will be examined first in a very general sense, followed by a more specific explanation of its meaning, in the context of measurement and experimentation, within the empirical-analytical paradigm. This particular view will then be contrasted against the implications held for the term in hermeneutic, phenomenological, anthropological, sociological and artistic research





paradigms.

Because of its intimate relationship with the concept of knowledge, this examination of "validity" will demonstrate:

1. That different kinds of knowledge are sought in different research paradigms.
2. That there are differences among skills involved in gaining each type of knowledge.
3. That each type of knowledge is tested according to relevant conceptions of epistemology which are operationalized within procedural rules for the conduct of inquiry.

#### Overview of the Concept

The fugitive nature of the concept "validity" has served to divide the ranks of knowledge seekers for centuries. Roots of the difficulty with the term lie in its fundamental relationship with the perennially indeterminate notions of "truth" and "reality". Valid knowledge should correspond with truth, and valid observations should correspond with reality, but the nature of these terms has remained evanescent despite centuries of philosophic endeavour.

Various schools of philosophy have adopted ontological positions (the first principles of which are untestable) for the purpose of developing theories to explain the nature of our world. These schools have



subsequently spawned multiform research traditions, some of which have proved more useful than others. Most have manifested themselves, in some way or other, in the field of educational research.

The concepts of truth and reality held by each school possess slightly or grossly different meanings depending upon the ontological base from which the logic of the tradition is constructed. Validity within each tradition is determined by observance of procedural rules that represent operational definitions of the epistemology characteristic of that tradition. The acceptability of these procedural rules is determined by the community of scholars that subscribe to each tradition.

Within educational research the question of validity appears to be significant across two domains of knowledge, the first relating to theories, and the second relating to observation (or measurement) of phenomena. These domains share an intimate and coterminous relationship as theory construction is dependent upon "valid" observations.

### Theory Validation

Kaplan (1964) has made somewhat global and pragmatic observations about the meaning of validity in both the above domains. With respect to theory validation, he suggested that at any given time a particular theory would be accepted by some scientists and not by others. Furthermore,



the theory may or may not be regarded as useful or appropriate in all possible cases of application by those scientists who accept it. In Kaplan's words,

The validation of a theory is not the act of granting an imprimatur but the act of deciding that a theory is worth being published, taught, and above all, applied - worth being acted on in contexts of inquiry or of other action. The acceptability of a theory will in any case be a matter of degree - more or less weight will be assigned to it, and it will always have a more or less limited range of justified application. (1964, p. 312)

Scientists, consequently, must exercise "good judgement" in scientific decision making with respect to selection and application of specific theories. Kaplan cautioned that a theory is not validated merely because it is accepted, but rather it is accepted because it is believed (by scientists) to be validated. Thus "good judgement" requires exercise of clearly defined norms. The norms identified by Kaplan as significant are i) the norm of correspondence, ii) the norm of coherence and iii) pragmatic norms.

Kaplan viewed the first of these as basic. The norm of correspondence is governed by the reality principle and asks the question "Does the theory fit the facts?" The difficulty implicit within this approach has to do essentially with the researcher's position regarding the "facts". Conduct of any inquiry will, of course, depend for its ultimate form upon the investigator's ontological stance. Kaplan suggested that what counts in the validation of any theory, with respect to "fitting



the facts" is convergence of the data that are brought to bear on it. This is a reasonable and pragmatic approach; however, it should be recognised that the data can only fit the "facts" as "facts" are construed within the ontological framework employed to conduct the research.

The norm of coherence was defined by Kaplan in three parts. The first was the degree of relatedness a specific theory is able to establish with the body of theory already instituted. He warned, however, that integration is a conservative notion which may serve to suppress scientific revolution, and that unyielding insistence for every new theory to fit those theories already established is characteristic of closed systems of thought. Kaplan went on to say,

What must in any case be taken into account in assessing a theory is the set of alternatives to it in conjunction with which it is being considered. That a theory is validated does not mean that it is probable, in some appropriate sense, but only that it is more probable than the other possible explanations.  
(p. 315)

This notion has important implications for the paradigm of aesthetic criticism.

Kaplan also elaborated upon two other conceptions of the norm of coherence - the norm of simplicity and the aesthetic norm. Although he recommended each should be sought, neither contains intrinsic justification. Theories should be no more complex than necessary; however, some explanations of nature require considerable complexity, and





a theory is not necessarily the most appropriate because it is the simplest. In addition, a theory that demonstrates intrinsic beauty (as in the execution of a mathematical proof, for example) does not have to be true - beauty is not necessarily truth. However Kaplan suggested that the aesthetic qualities of a theory have often played an important part in the context of discovery.

The third norm identified by Kaplan was the Pragmatic norm. The questions posed by this principle are simply "How effectively does the theory function?", and "How well does it serve our scientific purposes?" Kaplan cautioned that in terms of vulgar Pragmatism, the success or failure of a theory did not furnish necessary and sufficient information to validate or prove invalidity of a theory.

#### Validity in the Empirical - Analytical Paradigm

At this point it is necessary to return to the observation made earlier that various schools of philosophy, founded upon different ontological assumptions, have developed research traditions with varying degrees of influence on educational enquiry. By far the most significant of these, in education, during this century has been the empirical - analytical tradition, derived from the natural sciences.

Matson (1964) traced the origins of this tradition to Newton's conception of celestial mechanics. He



emphasised that methods of research in the natural sciences owed to the Newtonian scheme a notion of reality that permitted perfect separation of the observer and the observed, and the belief that all elements in the observable world were united in a mechanistic system of causal relationships. Referring to this idea in the context of atomic theory Matson said:

The evolution of every isolated mechanical system is rigorously determined by its initial state, and, more specifically, that exact prediction of its future behaviour follows directly upon knowledge of its present position and velocity. Originating with Newton's celestial mechanics, this rigorously deterministic doctrine had soon been extended to all fields of the physical sciences (and by eager extrapolation, to the social sciences as well). (p. 140)

In education, positivistic, behaviouristic and operationist approaches to research can be regarded as the direct consequences of the doctrine of mechanistic determinism. Bonner (1965) severely criticised those approaches, primarily because they removed the human component from scientific descriptions of the world.

Measurement. Given the ontological assumption of a world "out there" separate and distinct from the observer, the overarching criterion for validity of observation (and theories supported by those observations) within such traditions is most likely to be objectivity. The omnificent tool of scientists involved in this mode of inquiry is measurement. A highly specific set of procedural rules has consequently become established



in the social sciences for the purpose of ensuring validity of measurement. Concomitant with the evolution of these rules has come a differentiation of the notion of validity into two interrelated dimensions. The first deals with the definitional aspects of measurement while the second deals with predictive considerations (Kaplan, 1964).

Kaplan observed that both dimensions address themselves to the basic question "Does the instrument measure what it purports to measure?" The definitional aspect has to do with the meaning of the term naming the magnitude in question. This normally requires development of an operational definition for that magnitude. The predictive dimension is related to the empirical connections demonstrated by the test. (For example, how well an aptitude test will enable prediction of a particular set of behaviours in a job context.) Both dimensions are usually employed in determination of test validity.

Much controversy has raged among psychometricians about procedures for test validation that may be regarded as sound. In 1950 Cureton wrote Validity Reliability and Baloney in which he criticised current methods of computing validity coefficients. In 1954 the American Psychological Association set up a committee to stabilize definitions. Cattell (1964) criticised the work of this committee, claiming it succeeded too well in stabilizing too soon a relatively naive stage of psychological opinion. In addition



Cattell observed:

The way in which these parameters (for validity and reliability) were defined twenty years ago and the present habitual usages which derive from those definitions are now increasingly recognised to be misleading and provocative of inconsistent conclusions. This inadequacy and confusion becomes especially evident as testing extends from ability and achievement areas to the more subtle concepts required in personality and motivation measurement. (p. 1)

Cattell proposed that the validity of a test be considered along three independent parameters (or dimensions) producing eight types of coefficients of validity, four of which he considered to be particularly significant. The dimensions he identified were conceived to exist along the bi-polar continua of i) direct to circumstantial, ii) concrete to conceptual, and iii) natural to artifactual. The significant coefficients were regarded as i) concrete - direct, ii) conceptual - direct, iii) conceptual - circumstantial and iv) concrete - circumstantial. He suggested that all existing notions of validity could be accommodated under these categories. Also, he claimed certain popular existing uses of validity were either unfruitful or superfluous. Cattell's view was that "construct" validity and terms such as "face", "content", "predictive", "concurrent", "fiat", and "semantic" validities were either not related to validity at all or were more accurately conceived of as "utility" coefficients.

The meaning of Cattell's "dimensions" and "coefficients" will not be elaborated here, nor will his reasons for denigrating some popular conceptions of





validity be entered into. Rather, the purpose of reporting his ideas briefly is to indicate, in small measure, some of the operational difficulties encountered by measurement technologists conducting research within the empirical - analytical mode.

"Construct" validity is probably the most controversial aspect of the whole notion of validity in this paradigm. Messick (1974) claimed that construct validation is the process of gathering evidence, such as theoretically relevant empirical relations, to support the inference that an observed response consistency has a particular meaning. Messick criticised educational researchers for down-playing the significance of construct validity and claimed,

... all measurement should be construct - referenced. A measure estimates how much of something an individual displays or possesses. The basic question is, "What is the nature of the something?" It may be answered by referring to evidence in support of particular attributes, processes, or traits construed to underlie and determine task performance. (p. 8)

Handley (1973) questioned the classical process of determining construct validity, originally proposed by Loevinger (1957). Loevinger suggested that to validate a test each of "the" three approaches to validation (she called them components) should be employed sequentially. These were i) the substantive component, (formulation and judgement of a sound rationale for a construct), ii) the structural component (which determines the internal consistency of the scale), and iii) the external component



(related to appropriately established criterion validity).

Handley suggested that in practice this approach was seldom used in validating tests because of pragmatic questions such as shortages of time and money. In light of this he proposed that it may not be necessary to rigorously observe all three of these procedures.

The above discussion has served to indicate that certain difficulties with the notion of validity exist within the field of psychometrics, particularly with respect to construct validity. The source of the difficulty arises from the fact that a construct is not immediately observable in terms of a single, or simple set of behaviours. Collective (or sometimes singular) professional opinion is the only ultimate criterion for determining what, in fact, a construct measure is assessing. This is so because there is no other way to determine respectively i) if the "content" of the test is as it should be and ii) which other measures the test can reasonably be expected to correlate with, or predict.

Within the field of measurement two other related concepts are closely tied to the notion of validity. They are "reliability" and "error". Within the empirical - analytical model, as was pointed out before, the basic ontological assumption is of a world "out there" which can be objectively observed and measured. Validity in this context demands that measurements be relatively free from error. (Kaplan, 1964) Scientists recognise that no



measurement process is entirely free from some random fluctuation although the ultimate aim of measurement technology is to reduce error to an absolute minimum. Kaplan identified two sources of error - the first is inherent in the measuring instrument, the second is encountered when identical results are not obtained through repeated applications of the test in identical conditions.

The smaller random fluctuations become, the more reliable an instrument is said to be. Reliability is related to consensus in two ways: first, that different observers will ascribe the same measures to an identical situation; second, that the same observer will ascribe the same measures across situations that are believed to be the same. The primary quality demanded of a measure, to demonstrate reliability, is consistency. Reliability should not be confused with "accuracy", which is related to the assumptions about the magnitude being measured.

Kaplan's explanation of the meaning of validity and the function of the concept of error in valid measurement is illuminating and pragmatic.

The validity of measurement is often conceived in these terms (or rather, misconceived, as I shall urge): a valid measurement is thought of as one which is true to the actual magnitude being measured. I call this conception the fiction of the true measure. Metaphysically, it has both a realist and an idealist version. For the realist, facts are absolutely determinate in themselves, whatever may characterize our knowledge of them, so that objects and events are wholly definite as to quality and quantity. For the idealist, what is given in experience is always an approximation to an absolutely determinate abstract entity, which alone is



the proper objective of genuine knowledge. Both views, however, recognise that all actual measurements fall short of the determinateness of what they each conceive as the reality. It is for this reason that I speak of their conception as a fiction. What they call the "true measure" is what would result if we were to perform a measurement entirely free from error. But this is just what we cannot perform. .... From the standpoint of an empiricist, therefore, the "true measure" cannot have its meaning specified in these terms. For this specification amounts to explaining what we experience by something which lies outside experience. (p. 202)

Kaplan went on to suggest that a new measure is not necessarily more valid than an old one because it comes closer to the "real value" of the magnitude in question, rather it can be regarded as having greater validity if it is "more scientifically useful".

The thread of "consensus" runs consistently throughout efforts to determine validity and reliability of measurement. Consensus activity is employed to determine all types of validities mentioned above, be it agreement among judges or selected tests. Test reliability is contingent upon consensus of results across multiple applications or multiple administrators. The logic of the consensus model lies in the assumption that the higher the degree of agreement across the largest number of cases the closer to "objective" truth one approaches. And the closer one approaches to the "truth" or "reality" the greater the degree of validity of the measure used.

Experimental Design. The rules designed to enable achievement of valid observations and measurements within the empirical





analytical tradition extend beyond instrumentation. The circumstances surrounding administration of measurement must also be carefully controlled in order that the results of measurement describe only the magnitude in question and not the effects of other "contaminating" factors. These affect reliability directly and validity only as it encompasses reliability. The rules, designed to ensure appropriate conditions for measurement, are most explicit in the context of experimental design which is a procedure, borrowed from the natural sciences, that has been used extensively in education since Thorndike's time.

Campbell and Stanley (1963) in their widely used research handbook, have identified twelve classes of extraneous variables that can jeopardize the validity of experimental designs. These variables affect two forms of validity: internal and external. Internal validity has to do with the minimum conditions necessary for the results of a piece of research to be interpretable. External validity asks the question of generalizability.

An example of a question relating to internal validity could be "Did the (obtained) scores reflect the 'true' effect of the experimental variable, or was bias introduced by the maturation of experimental subjects?" The key assumption underpinning the notion of internal validity is that the magnitude in question exists independent from the observer. In other words, the magnitude (or observed response consistency) is not a product of the observer's "world view".



External validity asks the question of generalizability. Application of knowledge gained from a small "representative" sample to a wider, carefully defined population, is one of the central constitutive interests of the empirical-analytical mode of enquiry.

In order that internal and external validities are preserved, the conditions of research must be carefully constructed and controlled. Campbell and Stanley presented sixteen research designs that were considered to be more or less effective to this end. Numerous other researchers have elaborated upon this topic. Campbell and Stanley recommended a dozen writers between 1923 and 1962 who had done so.

### The Meaning of Validity within Alternative Research Paradigms

In recent years alternative research paradigms have begun to impinge upon the domain of education, a field traditionally anchored in the empirical analytical mode. Dissatisfaction has been voiced, from several quarters, with the underlying premises of quantitative research technologies. Among the most audible of the dissenters have been individuals from the ranks of art educators (as was reported in Chapter I).

The roots of this discontent have been illuminated by Matson's observation that, "the focus of scientific explanation has shifted from the 'thing itself', from



nature as an independent reality 'out there', to man's own observation of nature". (p. 144) He made this statement in light of Heisenberg's principle of uncertainty and Bohr's principle of complementarity.

These two principles precipitated the total eclipse of mechanistic determinism in the natural sciences. Notions of uncertainty and complementarity had their origins in the field of quantum physics, the most exact of the natural sciences. In brief, the principle of uncertainty declared that "the kind of information assumed by classical physicists as the prerequisite to exact prediction - i.e. the simultaneous knowledge of position and velocity (of particles) - was impossible of attainment in micro-physics". (Matson, p. 141) The procedures used to measure one variable affected the other such that the greater degree of accuracy obtained in one created a higher degree of uncertainty in defining the other. The degree of uncertainty obtained was governed by an irreducible minimum. The epistemological implication of this principle is that man cannot observe the course of nature without interfering with it.

The law of complementarity, on the other hand, is related to the inability of atomic physicists to choose between two theories (wave and particle) to explain the ultimate nature of matter. Both formulations have limitations, but each succeeds in explanation where the



other fails. However, each is mutually exclusive in that both explanatory pictures cannot exist simultaneously. The ultimate outcome of this situation is that each theory is accepted for the specific limited purposes at hand. The significant epistemological implication of this law resides in its demand for tolerance of ambiguity.

From his analysis of the history of the natural sciences since Newton, and the effect of extrapolation of its assumptions to the social sciences, Matson drew several conclusions.

His first general observation was prompted by Heisenberg's principle of uncertainty. Because man is unable to observe the course of nature without interfering with it Matson said,

The modern (or post-modern) image of the scientist as actor, as "participant-observer", rather than detached spectator, has led to renewed consideration of the process of observation as a form of interaction or "transaction" - terms which stress that contributions are being made from both sides of the measuring apparatus (p. 145).

The principle of complementarity led Matson to make two further observations that have particular significance in the context of phenomenological, hermeneutic and artistic research methodologies. Matson suggested that in light of the complementarity principle we must choose which segment of the whole we wish to understand, but the whole itself is beyond reach. He then observed that the behavioural mechanists' rigorous investigation of the subject matter of biology and psychology, in physiochemical





terms, toward the twin ends of prediction and control, was not the only possible aim of investigation in the human sciences. To understand, rather than to manipulate the subject matter, was a clear alternative. The understanding mode has achieved its most refined form in the field of phenomenology, as is demonstrated in the following section.

Polanyi (1951) pointed out that actions can be represented entirely in terms either of causes or reasons. This is a subtle distinction that rests upon acceptance or rejection of the assumption of free will accorded the human subjects of any investigation in the social sciences. Explanation of human behaviour in terms of causes presupposes the existence of causal factors outside the control of the subjects. This manner of explanation is related to the doctrine of mechanistic determinism and assumes "passivity" of the human beings in question. This assumption is critical to the empirical analytical mode because it permits prediction of future behaviour given similar conditions.

Explanation of action in terms of reasons, on the other hand, assumes free will and reasoning ability of subjects and has quite different implications for researchers in terms of interpretation of findings. Matson was impressed by this mode of inquiry as is indicated by this statement: "the organism is quite different from the machine and its living reactions possess an element of fundamental incalculability and unpredictability". (p. 160)



Polanyi suggested that both modes of inquiry could be used independently but acceptance of one mutually excludes the other. The two frames of reference, therefore, are complementary and relate directly to Bohr's principle in a social sciences context.

Validity assumes two distinctly different meanings within each mode. In the causal framework, validity of explanation is determined by the closeness with which explanations approach the external "truth" i.e. satisfactory description of the causes of behaviour. The rules governing validity are those outlined in the previous section of this chapter - consensus, and predictive ability of explanation are of paramount importance.

Within the understanding mode a valid explanation requires the investigator to satisfactorily comprehend and report the subject's own interpretation of his world. The criteria for "satisfactory comprehension and reporting are varied and complex. Within the understanding mode, various schools of philosophy (which are related because of their common constitutive interest in this basic position) have coped with particular aspects of explanation in different ways. Validity, accordingly, has slightly different meanings in each situation.

Explanation within the "understanding" framework is considerably more complex than its alternative. If one accepts the "human dimension" (i.e. reasoning ability and free will) of investigative subjects, and, in addition, acknowledges the inability of the researcher to observe



without affecting the natural course of events, one ineluctably includes the scientist in the phenomenon he is investigating. This peculiarly difficult circumstance may be considered analogous to the situation where one attempts to push a bus in which one is sitting. (Berger & Luckmann, 1966)

In this same framework, the notion of validity is correspondingly complex. For example, the researcher may see fit to explain his subject's conception of the world in terms of his own conception of that subject's explanation - which he must also make clear. Validity of explanation in this context becomes largely a problem of objectively describing a subjective world.

Each of the philosophic traditions discussed below accepts different positions with respect to the nature of "reality" and the means employed to "explain" it. Some controversy exists within various traditions on this issue. Van Manen (1974) observed that these alternate research orientations are based upon conceptions of objectivity and validity different from the empirical analytical approach. Gouldner (cited in Van Manen, 1974) reported the emergence of new criteria of objectivity in social science and social theory. Objectivity was not now regarded as "absolute", resting on observational empirical evidence, but arose from interpretation of the meaning of that evidence which leads to the construction of consensus (or truth) by means of convincing accounts, acceptable by the "intersubjective community" of scientists.



## The Hermeneutic Tradition

"Hermeneutic research" means interpretative inquiry, and has its historical roots in biblical exegesis, the interpretation of literature, and law. According to Habermas (1972) "Knowledge in the hermeneutic tradition is always mediated through an interpreter's pre-understanding." Understanding of meaning is directed in its very structure toward the attainment of possible consensus among actors in a framework of self understanding derived from tradition.

Validity in interpretation has been a problematical issue for many years. Hirsch (1967) claimed,

Since all humane studies, as Dilthey observed, are founded upon the interpretation of texts, valid interpretation is crucial to the validity of all subsequent inferences in those studies. The theoretical aim of a genuine discipline, scientific or humanistic, is the attainment of truth, and its practical aim is agreement that truth has probably been achieved. Thus the practical goal of every genuine discipline is consensus - the winning of firmly grounded agreement that one set of conclusions is more probable than others - and this is precisely the goal of valid interpretation". (p.p.viii-ix)

Hirsch demonstrated the problematical nature of validity by identifying three other interpretive theories - each of which held different conceptions of the "truth". The "pragmatic theory", most commonly employed by the legal profession, in simplistic terms implies that the law means what the judge says it means. The second theory, often used in biblical exegesis, claims that the bible holds a new revelation for each successive generation. The third (which) Hirsch called "Autonomism", maintains that literary texts belong to a distinct ontological realm where





meaning is independent of authorial will. Hirsch observed that this is a mystical idea that has never been persuasively defended. He also held that all three of the above views implicitly deny the possibility of validity in an absolute or normative sense.

For Hirsch all valid interpretation is founded on the recognition of what the author meant, and validation is the process of showing, in a particular case, that the author's verbal meaning has probably been achieved. The crucial problem in judging between disparate interpretations is usually the comparative weighting of relevant evidence.

The issue of authorial will, or intentionality, has been a moot point in philosophical circles for centuries. To accept authorial will as the criterion for validity is akin to adopting the empirical analytical position of a reality (author's intention) "out there" separate from the interpreter who objectively (by consensus) identifies that will. Hirsch's position in hermeneutic research is unusual. He himself recognised that the possibility of an absolutely valid interpretation has been regarded with skepticism in recent years.

Recent uses of hermeneutic methods have tended to stress "sense-making", acquisition of insights into human experiences and sharing of common meanings. Van Manen (1974) said,

Interpretive science has a research-guiding interest to clarify, authenticate, uncover, or to bring the meaning structures expressed by the productive forces of the human cultural process into full human awareness.



....The inherent telos is understanding (Verstehen) aspects of the human life world, in the sense of gaining insights into the processes and results (objectivations) of human cultural activity in the way of text, of text analogues, expressing life projects, sense-making and interpretive practices, Human actions and intentions, and the way in which man meaningfully experiences, and emotionally and intellectually appropriates the world. (p.11)

This approach emphasises the actor (interpreter), rather than the author, and the focus is upon "meaningful understanding" of the text. Intentionality is not as significant, for valid interpretation, as construction of intersubjectively shared meanings. The criterion for validity becomes, in this instance, referential adequacy of the constructed meanings. (Both intentionality and referential adequacy are discussed in relation to the artistic paradigm which shares common goals with hermeneutics.)

In an educational context Mann (1969), and later Willis (1975) proposed that "curriculum" can be considered as a literary object and hence may be treated, by an interpreter (critic), in the same way a literary critic deals with a literary work. Addressing himself to what is essentially a hermeneutic problem, Mann suggested that criticism should focus on the curriculum itself, rather than on the author. However, the critic's statements are not necessarily logically derived directly from the data before him (the curriculum) but are the artful result of his adoption of a methodology that allows expression of his own critical principles, assumptions, and choices about subject matter. Mann did not deal directly with



intentionality, although the preceding observation serves to imply that he did not consider validity of interpretation important in the normative sense.

Mann did address the problem of objective and subjective knowing (Polanyi, 1964). He suggested the common tendency to associate the former with science and the latter with art was at least an inadequate conception, if not plain wrong. At every point of inquiry, from selection of problems to his drawing of conclusions, the scientist's work is a combination of rule-governed procedures and "heuristic leaps" beyond those rules. His major point was that the "personal" component of knowledge is all-pervasive in the activities of artists and scientists.

Following from this observation, Mann then suggested that the "aesthetically oriented" critique and the scientifically conceived research programme are less different than may be imagined. The research project begins with a decision about what to select for attention and that decision is governed by what one's personal knowledge leads him to believe will be valuable and fruitful. The curriculum critic's problem (like any interpreter's) is to select from an inexhaustible realm of designs and meanings those he will attend to, and that decision is grounded in personal knowledge. Mann emphasised that this does not mean that the content of what one will disclose in a critique is fixed in advance by prior personal knowledge but it does mean that the dimensions, in



terms of which content will be sought, are fixed. In other words, the critic approaches the phenomena to be examined with a set of predispositions in the form of highly abstract models of what designs it would be of value to discover. These models do not determine what meanings will be disclosed in a curriculum, but they suggest the ethical dimensions which, if found in a design, will be regarded as worth attending to.

Validity in Mann's conception of interpretation is almost a "non-issue". He does not suggest that there is an incontrovertible "truth" to be revealed by interpretation. Rather the discovery of new meanings, through drawing upon inwardly intuitively held unformalized knowledge, is paramount.

Three differing views of hermeneutics have been presented above. The first (Hirsch) required agreement that the author's intended meaning has probably been achieved for a valid interpretation to occur. The second (van Manen) emphasised construction of intersubjectively shared meanings, an approach that recognises "meaningful understanding" for the actors involved. Validity relies, in this instance, upon the referential adequacy of ascribed meanings. The third view (Mann and Willis) placed ultimate significance upon discovery of "new meanings" grounded in the critics' "personal" knowledge. Validity, in the sense of generating a normative (objective) truth, is not a significant issue within this framework. Pragmatic value (appropriateness) of the interpretation for the problem





at hand, is the criterion determining the strength (or validity) of the interpretation.

### The Phenomenological Tradition

One of the most significant writers in the field of phenomenology is Alfred Schutz. Explanation of phenomenology as Schutz envisaged it, however, is beyond the scope of this chapter. To serve the purpose at hand, several of the central ideas of this tradition are presented to encapsulate the meaning of validity in a phenomenological research context.

Phenomenology is an interpretive science, as is hermeneutics. Where hermeneutics is concerned primarily with the interpretation of texts, or text analogues, phenomenology is directed towards constructing meanings for the "here and now" world of everyday social life. The difference is primarily one of time - hermeneutics is concerned with analysis of past events, whereas phenomenology interprets the present.

The important assumption upon which all phenomenological theory is constructed is that "each one of us is part of an ongoing world of everyday affairs which is, for the most part, taken for granted in its essential being" (Natanson, 1973). The "common sense" world is the scene of man's actions conducted, not so much for the purpose of understanding it, but to effect changes within it. The structures of daily life in the common sense world are typically taken for granted. Each



individual locates himself in daily life in a particular manner which Schutz (1973) called his "biographical situation". The essential significance of this construct is that the formative period of each life is realized in a unique way. Although common sense reality is afforded all men in terms of a universal validity, the way in which its forms are translated into an individual life will depend upon the totality of the experience a person builds up in the course of his concrete existence.

The central assumption, for researchers in the phenomenological tradition, is that it is the meaning of our experiences, and not the ontological structure of objects, which constitutes reality. The way in which individuals ascribe meaning varies from case to case. Schutz developed a theory of "multiple realities" which was designed to explain the reason for differences between individual views of the world. He expressed it as follows,

There is an intrinsic order of our perceptions of outer objects, or of so called inner experiences of fantasms, and even of dreams, which separates them from all other realms and constitutes them, according to one formulation, as separate provinces of meaning. Here again we have within limits the freedom to select one of these realms as our system of reference, that is to "live in one of these orders or, to bestow upon one of them the accent of reality. We have, then, several concurrent and competing orders of reality - that of our everyday life, that of the world of our fantasy, of art, of science etc., among which the first is paramount, because only within it is communication possible. (p. 339)

Within the world of everyday life Schutz identified additional constructs which are important for consideration of the development of meaning structures by the individual.



Significant among these is the notion of intersubjectivity. One aspect of intersubjectivity important for understanding the validity of scientific observation is reciprocity of perspectives. Within the natural attitude of common sense thinking in daily life it is taken for granted that other intelligent human beings exist. It is assumed that the objects of the world are accessible to their intelligence. It is also taken for granted that, strictly speaking, the "same" object must mean something different to individuals sharing the common-sense world because of i) differences of viewpoint (time and/or position) and ii) differences in biographically determined situations (which affect purposes at hand and respective systems of relevancies).

Common sense thinking overcomes these differences by two idealizations i) the idealization of interchangeability of standpoints (that the object should have the same meaning if I change places with my fellow man) and ii) the idealization of the congruency of the system of relevancies (that differences in perspective attributable to biographical situation are irrelevant for the purposes at hand until counter evidence becomes available).

The difference between the research interests of the natural scientist and the social scientist (in the phenomenological mode) was expressed by Schutz as follows,

The world of nature, as explored by the natural scientist, does not "mean" anything to



molecules, atoms and electrons. But the observational field of the social scientist - social reality - has a specific meaning, and relevance structure for the human beings living, acting, and thinking within it. By a series of common sense constructs they have pre-selected and pre-interpreted this world which they experience as the reality of their daily lives. It is these thought objects of theirs which determine their behaviour by motivating it. The thought objects constructed by the social scientist, in order to grasp this social reality, have to be founded upon the thought objects constructed by the common sense thinking of men, living their daily life within their social world. (p. 59)

This view clearly echoes Polanyi's distinction between research designed to discover causes, and that seeking reasons for human action.

The difficulty for the social scientist in the phenomenological mode is to develop a method to deal, in an objective way, with the subjective meaning of human action. To do this Schutz proposed that the social scientist detach himself from his biographical situation within the everyday world and adopt a specific "scientific attitude". The nature of the scientific problem he defines for himself re-structures his system of relevancies and taken-for-granted courses of action in the everyday world. He assumes the attitude of a "disinterested" observer in the social context. The problem becomes the "locus" of all possible constructs relevant to its solution.

The scientist must adopt what is considered by his fellow scientist as knowledge, or "show cause" why he is unable to do so.

To explain the subjective meaning of human action





the scientist develops model constructs of the social world. Schutz claimed the models can come to terms with social reality only if they are developed according to the following three postulates:

i) The postulate of logical consistency.

The system of constructs designed by the scientist must be clearly articulated and compatible with the principles of formal logic.

ii) The postulate of subjective interpretation.

The model constructed by the scientist must be formed so that the observed facts are explainable in terms of the typical activities of an individual mind.

iii) The postulate of adequacy.

Each component of the scientific model of human action must be so presented that any indication (by the model) of action performed in the life-world must be understandable for the actor himself. It should also be understandable for his fellow men in terms of common-sense interpretation of everyday life.

Validity of interpretation in the phenomenological mode of inquiry is virtually defined by these postulates. "Scientific" explanation of human action must be logical (for internal consistency), representative of the "subjective" activity of a human mind and understandable by the actor and his fellow men.

Within the community of scientists the notion of intersubjectivity is important for communication of



knowledge gained through research. The idealizations of interchangeability of standpoints and congruency of the system of relevancies perform a significant role in this respect.

### The Sociological and Anthropological Traditions

The research paradigms mentioned above demonstrate a full range of possibilities with respect to conceptions of reality, objectivity and (subsequently) validity, from logical positivism on the one hand to the subjectivist focus of phenomenology on the other. Within the fields of Sociology and Anthropology various "spin off" methodologies have drawn from selected positions between the two extremes.

Closely aligned to the phenomenological mode of inquiry are found linguistic and ethnomethodological approaches (Cicourel, 1972), and sociology of knowledge explanations of the world (Berger and Luckman, 1966).

Other methodologies, although demonstrating research procedures that claim variance from empirical analytical approaches, serve the ends of statistical research. For example, Smith and Geoffrey's (1968) Complexities of an Urban Classroom was a study that claimed an ethnographic approach yet served the ends of empirical analytical science (causal) explanation, prediction and control. Pohland's (1972) participant observation study employed (initially) a holistic, unstructured, "understanding" mode of inquiry, yet maintained the



empirical - analytical test of validity (consensus of observations through multiple instrument, multiple person, multiple occasion observations).

A complete analysis of the "mutant methodologies" that have flourished in education upon the nutrition of alternate philosophies will not be undertaken here. Instead it is sufficient to report their existence and to note that it is the conception of validity (of knowledge), held by these methodologies, that indicates their philosophic allegiance, ontological position and ultimate purposes, rather than the research methods they boast.

### The Artistic Tradition

The epistemology of art is based upon theories of interpretation and, because of this, may be regarded as a branch of the hermeneutic sciences. Artistic knowledge is gained from interpretation of visual configurations and relies upon mediation by an interpreter (who possesses a wide background of immediately relevant knowledge) to clarify meaning.

Interpretation of the meaning of artistic objects is complicated, for the following reasons:

1. The concept of art is open ended. There is no necessary and sufficient set of conditions for the judgement of artistic quality (or validity) (Weitz, 1966).
2. Artistic knowledge differs from symbolic and



scientific knowledge because of its particularity.  
Essentially every aesthetic object is incomparable.  
(Phenix, 1964)

3. A work of art does not necessarily attempt  
to approximate "reality" or "truth" (Beardsley, 1958;  
Dearden, 1968).

These interrelated reasons have the ultimate  
effect of obfuscating the conditions necessary for  
determination of "valid" aesthetic knowledge. Dearden  
(1971) put these reasons together nicely in the following  
discussion of validity of aesthetic judgement:

The validation procedures of aesthetic judgement  
present considerable difficulty to anyone seeking  
to state them, partly because there is no obvious  
sense in which the work of art is an attempt to  
approximate to something already "given". There  
appears to be no reality about which the work  
of art is an attempt to state the truth, as in  
science and history....In science and history,  
there is an attempt to find laws and furnish  
narratives which in some sense "correspond" to  
what actually occurs or has occurred; otherwise  
they would be fictions and fantasies adrift from  
reality.

But a work of art is not already implied,  
waiting only to be traced, or "there", waiting  
to be discovered. It is the artist's creation  
and, in a sense, its own world. (p. 296)

....Yet he (the critic) cannot provide us  
with any set of simple rules for the process  
of criticism, because a feature which is entirely  
fitting in one work would be painfully dissonant  
or unbalancing in another. (p. 297)

The three conceptions of validity of interpretation,  
discussed in the context of the hermeneutic tradition,  
are reflected in the positions adopted by aestheticians  
with respect to valid aesthetic judgements. Briefly,  
these are:





1. Statements about an aesthetic object are valid if they correspond with the artist's intention.
2. Statements about an aesthetic object are valid if they correspond with the meaning, intention, or truth implicit within the work.
3. Statements about an aesthetic object are valid if they correspond with the meaning ascribed to the object by the viewer as a result of his interaction with it.

Each of these criteria deserves lengthy elaboration which is not possible here. Instead, the implications (for validity) and practical value of each position will be briefly assessed.

Intentionality. The first argument for validity assumes that because works of art are the products of deliberate human activity they are largely as they were intended to be by the individuals who made them. In order to understand the "true" meaning of the work the viewer must "see" the artist's intention.

Intentionality can refer to the artist's desired result (artistic or otherwise) or his intended meaning. The "desired result" refers to goals as broad as i) expression of the artist's initial conception of the final configuration (held in his head), ii) gaining of a specific effect known about in advance and iii) achievement of non-artistic results (such as fame).



The "intended meaning of the artist" is subtly different from the above and could refer to 1) the "message" built into the work or 2) specific meanings of particular symbols within the work (Kuhns, 1966). Other more tenuous arguments for intentionality refer to the "psychological state" of the artist.

The position that takes the intention of the artist as the final court of appeal for validity in interpretation is not a strong one. Beardsley (1958) discounted this view by suggesting that a general principle of philosophy should be kept firmly in mind in assessing a work of art.

If two things are distinct, that is, if they are indeed two, and not one thing under two names (like the Vice President of the United States and the Presiding Officer of the Senate), then the evidence for the existence and nature of one cannot be exactly the same as the evidence for existence and nature of the other. (p. 19)

The nature of the artist's intention is only indirect evidence of the nature of the object itself - it is external to the object. There is no problem in interpretation when internal and external evidence agree. The difficulty arises when they conflict e.g. if the artist says the work represents one thing whereas the senses indicate something quite different. One must choose between two "realities" - the object in the artist's mind, or the public object visible to the senses. To accept the former, Beardsley claimed, can lead to the wildest absurdity - an artist can make his art object symbolize anything he wants simply by saying it does.



Implicit Meaning. The second position, that meaning, intention, or truth exists within works of art (independent of the artist's intention or viewer's frame of reference) is a most contentious view. The reader will recall that Hirsch referred to this type of theory as "Autonomism". Philosophic controversy has given rise to a large number of theories in this area of aesthetics.

Kuhns indicated three possibilities of "intention" in works of art. The first had to do with the suggestion that the elements of a work of art cohere as an organic whole in such a way that "purposiveness" is apparent. He referred to Kant's view that a work of art is like a natural object - it seems to be composed for some (particular) end.

The second view, influenced by Psychology, holds that a work of art, because of the way in which the elements are arranged, will cause the beholder to respond in a particular fashion. The parts of the work exist for the specific end to which they contribute. This end is a demand for a particular response from the viewer.

The third possibility of intention in art works, identified by Kuhns, has to do with "supra personal" purposes. An artist allows the forces of autonomous creativity to work through him so that specific meaning is manifested in the work by an apparently "foreign impulse" not under the artist's control.

Beardsley (1958) examined the claims of



"proposition" and "revelation - intuition" theories that "truth" exists in works of art. "Propositions" are statements claimed to have an existence independent of the vehicle used for their expression. Proposition theorists claim that these statements are implicit within works of art.

The "revelation" or "intuitionist" theory claims that the true meaning of a work of art is intuited by the beholder. Verbalization of the knowledge gained, however, will distort its meaning.

All of these (and related) theories contain enormous difficulties arising from their claims for independent existence of "meaning", "truth", or "intention". Such a position is essentially metaphysical, which consequently limits its pragmatic value - the criteria for validity are not testable.

Ascribed Meaning. Probably the most commonly used theory to explain artistic knowledge is that meaning results from the viewer's experience of an art object. Dewey's (1934) classic book Art as an Experience is a testament to this point of view.

The question of validity of knowledge gained through "personal experience", is usually determined by referencing knowledge against the object. In other words, conclusions about the work must be supported by the senses. Philosophers who propound this point of view generally recognise that individuals ascribe various meanings to





the same art object because of their "biographical location" (to borrow Schutz' term). The question of consensus (in determining the "most valid" knowledge), however, is less clear-cut.

Phenomenologists, for example, regard all individuals' interpretation of their own experiences (artistic or otherwise) as equally valid - the knowledge of the critic is as valid as that of the neophyte.

Others, such as Ecker and Kaelin (1972) suggest that knowledge, to be valid, must represent a "relevant response" to the public art object. This implication is obvious in the following quotes:

Even the artist's reaction to his own work need be no different from that of any other viewer who knows what he is looking at. What makes it possible for communication to take place at all is the existence of a sharable, public object whose very structures control all relevant responses to it. (p. 263)

All language used to describe the public objects of human experience must be grounded in ostensive definitions - those which point to their objects. (p. 264)

The latter quotation represents part of an argument for "referential adequacy" of observations (knowledge) about art objects. It is a useful argument in that artistic knowledge can be justified by a process of i) defining terms used, then ii) proceeding logically from the work to supported conclusions or judgements. Where the consensus model is used (as in Habermas' [1971] conception of hermeneutics) validity is determined by agreement in choosing the most strongly supported



argument from among the competing alternatives. This model is also tolerant of ambiguity in that two conflicting, but equally well supported arguments, may be accepted as equivalent in validity.

The ultimate aim of hermeneutics, however, is consensus. The mediator (usually a critic) plays a central role in directing attention, and "re-educating" the perception of actors in the social context. The critic in the artistic paradigm brings to bear widely differentiated knowledge of artistic phenomena. By use of descriptive language he directs the beholder's attention to specific designs or aspects of the work he considers to be relevant. Ecker and Kaelin observed that:

Any person capable of understanding the critic's language may, by consulting his own experience, just as readily focus upon those features of the public object referred to in the artists use of language. (p. 266)

However, the temptation for common attitudes and shared meanings of critical discourse to be elevated to an arbitrarily restrictive elitist culture should be strongly resisted, the authors warned.

The difficulties implicit within the critic's role, with respect to validity, are the same as those indicated in Mann's argument for curriculum criticism discussed earlier.

Although the critic may defend the validity of his conclusions on the basis of referential adequacy he must initially select, from an infinite universe of



designs, those which he considers important. To do this (as Mann suggests) he draws upon his "personal" knowledge of designs that would be of value to discover. He also brings to bear his own paradigmatic criteria for the making of judgements about the work. (Hence it is possible for two critics to arrive at different conclusions).

With respect to validity of the initial choice of variables selected by the individual critic for attention, no criteria may be employed. Consequently the critic's only defence for the validity of his choice is upon phenomenological grounds.

The notion of reliability deserves mention in connection with the artistic - hermeneutic paradigms. If it is accepted that two differing conclusions about the same work of art can be regarded as valid (because each is referentially adequate), then reliability is not a necessary condition of validity as it is in the empirical analytical paradigm.

#### Summary of Chapter IV

In this chapter the meaning of the term validity was explored for the purpose of providing the groundwork for explanation of the validity of the model developed in this study. General observations revealed that the root meaning of validity is "strength", so valid knowledge may be regarded as that which provides the most effective means to given epistemological ends. These ends usually



were to discover "truth" or "reality", conceptions of which varied within alternate research paradigms. In a broad sense, theories are validated according to three general norms - the norm of correspondence, the norm of coherence, and pragmatic norms.

Within the empirical - analytical paradigm, validity of knowledge rests upon the assumption of objectivity (separation of subject and object). Objectivity is determined by consensual procedures. The primary tool of the empirical - analytical researcher is measurement. Two components are required to establish validity of a measuring instrument - the definitional component and its empirical connections. Controversy was reported in the field with respect to procedures used to determine the definitional aspects of validity.

The experimental design was found to be significant in this tradition. Two forms of validity (internal and external) were observed as necessary for useful interpretation and generalization of findings. The central interest of empirical - analytical research is discovery of causal relationships for the twin ends of prediction and control.

Alternative research traditions generally were found to seek explanation of human action in terms of reasons rather than causes. The research interest of these paradigms generally is "understanding" rather than prediction. In addition these paradigms were generally more tolerant of ambiguity. Within the hermeneutic tradition of interpretation of texts and text analogues,





three different positions were reported with respect to criteria held for determination of validity. They were correspondence with the author's intention, construction of intersubjectively shared meanings (sense making) and the discovery of new meanings. Within the phenomenological paradigm, to be valid, explanations of human actions should be understandable (and recognised as correct interpretations) by the actor, and his fellow man, in the commonsense world of everyday life. Various "spin off" methodologies have developed from hermeneutics and phenomenology, within anthropology and sociology, that demonstrate varying degrees of "purity". The identification of a criterion for validity is more likely to be achieved by studying the ends which the study served than the methodology employed.

Within the artistic paradigm, interpretation of the meaning of art objects is complicated because the concept of "art" is open ended, artistic knowledge is particular, and art does not attempt to approximate something else (truth or reality) which can be used as a criterion for validity. Three basic criteria for validity of artistic judgements were observed to be held by aestheticians. They were correspondence with the intention of the artist, correspondence with the meaning implicit within the work, and referential adequacy. The first two were found to be inadequate in pragmatic terms. The latter was found to be useful as a means of determining



"strength" of observations made about artistic phenomena.  
However, the initial choice (by the critic or neophyte)  
of which variables to attend to can be defended on  
phenomenological grounds only.



# REFERENCES

- Beardsley, M.C. Aesthetics: Problems in the Philosophy of Criticism. New York: Harcourt, Brace & World Inc., 1958.
- Berger, P.L., & Luckmann, T. The Social Construction of Reality: A Treatise in the Sociology of Knowledge. New York: Doubleday & Company Inc., 1966.
- Bonner, H. Two Views of Science. On Being Mindful of Man: Essay Toward a Proactive Psychology. Boston: Houghton Mifflin Co., 1965.
- Campbell D.T., & Stanley, J.C. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally College Publishing Company, 1963.
- Cattell, R.B. Validity and Reliability: A Proposed More Basic Set of Concepts. Journal of Educational Psychology, 1964, 55(1), 1-22.
- Cicourel, A.V. Cognitive Sociology: Language and Meaning in Social Interaction. London: Penguin Books Ltd., 1973.
- Cureton, E.E. Validity Reliability and Baloney. Educational and Psychological Measurement, 1950, 10, 94-96.
- Dearden, R.F. The Aesthetic Form of Understanding. In R.A. Smith (Ed.) Aesthetics and Problems of Education. Chicago: University of Illinois Press, 1971.
- Dewey, J. Art as Experience. New York: G.P. Putnam's Sons, 1934.
- Ecker, D.W. & Kaelin, E.F. The Limits of Aesthetic Inquiry: A Guide to Educational Research. In P.L.G. Thomas (Ed.) Philosophical Redirection of Educational Research. The Seventy-First Yearbook of the National Society for the Study of Education. Part I. Chicago, Illinois: University of Chicago Press, 1972.
- Habermas, J. Knowledge and Human Interests. Boston: Beacon Press, 1971.
- Handley, W.B. Substantive, Structural and External Components of Construct Validity in Instrumentation. Unpublished doctoral dissertation, University of Alberta, 1973.



- Hirsch, E.D., Jr. Validity in Interpretation. New Haven: Yale University Press, 1967.
- Kaplan, A. The Conduct of Inquiry: Methodology for Behavioral Science. Scranton, Pennsylvania: Chandler Publishing Company, 1964.
- Kuhns, R. Criticism and the Problem of Intention. In R.A. Smith (Ed.), Aesthetics and Criticism in Art Education. Chicago: Rand McNally & Company, 1966.
- Loevinger, J. Objective Tests as Instruments of Psychological Theory. Psychological Reports, 1957, 3, 635-694.
- Mann, J.S. Curriculum Criticism. Teachers' College Record, 1969, 71(1), 27-40.
- Matson, F.W. The Broken Image: Man Science and Society. New York: George Braziller, Inc., 1964.
- Messick, S. The Standard Problem: Meaning and Values in Measurement and Evaluation. Presidential Address to the Division of Evaluation and Measurement of the American Psychological Association at New Orleans, August 30, 1974.
- Natanson, M. Introduction. In M. Natanson (Ed.), Alfred Schutz: Collected Papers I: The Problem of Social Reality. The Hague: Martinus Nijhoff, 1973.
- Pohland, P. Participant Observation as a Research Methodology. Studies in Art Education, 1972, 13(3).
- Polanyi, M. The Logic of Liberty. Chicago: University of Chicago Press, 1951.
- Polanyi, M. Personal Knowledge. New York: Harper & Row, 1964.
- Smith, L.M., & Geoffrey, W. The Complexities of an Urban Classroom: An Analysis toward a General Theory of Teaching. Toronto: Holt, Rinehart, and Winston Inc., 1968.
- Schutz, A. Common Sense and Scientific Interpretation of Human Action. In M. Natanson (Ed.) Alfred Schutz: Collected Papers I: The Problem of Social Reality. The Hague: Martinus Nijhoff, 1973.
- van Manen, M. Some Search Orientations for Research in Social Studies Education. Paper presented at the NCSS Research Advisory Committee Session Chicago. Nov. 27, 1974.





Weitz, M. The Nature of Art. In D.W. Ecker & E.W. Eisner (Eds.), Readings in Art Education, Waltham, Mass.: Blaisdell, 1966, 49-56.

Willis, G.H. Curriculum Criticism and Literary Criticism. Journal of Curriculum Studies, 1975, 7(1), 3-17.



## CHAPTER V

### The Model

#### Introduction

This chapter is comprised of two parts. In the first section development of the model is described. This description includes; i) procedures followed to develop the model, ii) discussion of the nature of empirical and theoretical data that provided information for decisions about model structure, iii) a brief report of the pilot testing of the prototype of the model, iv) a critical analysis of the pilot test, v) the rationale that directed revision of the model.

The second section consists of an explanation of the revised model in terms of, i) its dimensions, ii) the relationships operating among those dimensions, iii) its methodological assumptions and iv) conceptions of validity appropriate to its implied research methodologies.

#### Part I: Development of the Model

##### Procedures Followed

The general procedures followed in developing the model were:



1. An analysis of two "live" curriculum development projects was conducted for the purpose of verifying significant operational components (of curriculum development) derived from the literature. These components were "Agents", "Content", "Implementation", "Outcomes", "Rationale", and "Facilities" (later expanded to become "Resources"). Verification was achieved by classifying topics discussed at development meetings of the two curriculum projects into the above six categories. This classification was conducted by a panel of six judges who demonstrated an average percentage of agreement of 76% across both projects. Further explanation of this procedure is discussed in Chapter VI.

2. Salient ideas that were compatible with the purposes of the study were selected from the literature to act as governing principles in construction of the model.

3. The theoretical and empirical dimensions were combined in a tentative model. (Appendix "A")

4. The model was pilot tested on one curriculum project.

5. Revisions to the model were undertaken on the basis of information derived from the pilot test.

### Description of Data

Data for construction of the model were drawn from empirical and theoretical sources. The empirical sources were two curriculum development projects that demonstrated very different kinds of characteristics. One was a short conference type programme conducted by the Fine Arts Council of the Alberta Teachers' Association. The other was a long



term curriculum development project conducted by the Edmonton Public School Board.

The 1975 Fine Arts Council Conference. The Fine Arts Council is a specialist sub-committee of the Alberta Teachers' Association. Each year a three day conference is held by this council, usually during the latter half of September. The conference attracts approximately two hundred members each year. Attendance in 1975 was divided roughly equally among practitioners in art, music, and drama. Participants included elementary and secondary teachers (specialists and non-specialists), administrators, tertiary instructors, Department of Education personnel, students, and executive members of the A.T.A. The conference programme included practical workshop sessions in the specialist areas of art, music, and drama, formal and informal discussion sessions, presentations of a philosophic nature, materials displays, and a variety of social events.

The Edmonton Public School Board Art Appreciation Project. This project was officially initiated with the submission of a proposal by the EPSB Art Supervisor to the Department of Research and Evaluation, Edmonton Public School Board, on March 29, 1974. The problem identified in that proposal related to insufficient guide materials in the Art Appreciation segment of the Alberta Elementary Art Curriculum Guide.

With the assistance of a committee of four elementary teachers and the art consultants, it was proposed that guide materials might be prepared to assist in





rectifying this problem. Initial guide materials were generated by two committee members during the summer of 1974. Refinement of these materials was undertaken during the winter of 1974-75, the summer of 1975, and the winter of 1975-76. Also during the summer of 1975 additional materials were generated to complete the first of a total of four sections planned to comprise the whole curriculum unit. Half-day committee meetings were held once per month during the school years 1974-75 and 1975-76 for the purpose of refining Section I of the materials.

In September 1975 a pilot study of the Section I materials began in 30 classrooms. Teachers using the materials were volunteers, selected on a first-come first-served basis.

Data were collected, relative to these projects, in the form of field notes made at the planning and development meetings of each committee. Focus of the observations made was upon the nature of the issues discussed. It was intended that these notes would serve as a basis for identification of the significant components comprising the curriculum development process.

Salient Ideas from the Literature. The theoretical sources of data were varied. In Chapter III "curriculum evaluation model" was described as a concatenation of intricate concepts that included, in general terms, theories of curriculum, curriculum development and



evaluation. Of these three it was noted that only curriculum development was readily observable in an empirical fashion. The other two were regarded as problematic abstract concepts.

The broad areas of the literature examined for relevant ideas were: general evaluation, evaluation of art, aesthetics, and criticism and research methodologies related to the artistic (interpretive) paradigm such as hermeneutics, phenomenology, and anthropology. Also covered within general evaluation and evaluation of art, were theories of curriculum development.

From this review several salient ideas were isolated. These ideas contributed towards determination of the final form of the model.

Within the field of general evaluation theory a survey of existing models of evaluation was conducted for the purpose of isolating common "components". These will be discussed in greater detail in the next section. In addition, three significant principles of evaluation were identified.

The first is appropriately related to an overall conception of "evaluation". Several writers have advanced relatively congruent definitions of "evaluation". Put together, these ideas suggest that evaluation is a process of defining, obtaining, and providing useful information about the worth of programmes, products, and procedures, to decision makers, to assist in their



judgement of decision alternatives. (Stufflebeam, Foley, Gephart, Guba, Hammond, Merriman, Provus, 1971; Stake, 1968; Worthen and Sanders, 1973).

The second significant idea from general evaluation theory was Apple's (1974) suggestion that evaluation is a process of social valuing and requires choosing, from a range of competing value systems, one that gives meaning to educational process or products. The taproot of this idea is anchored in the philosophy of interpretive sciences. Its major implication is that there are no "absolute" criteria against which observed educational events can be tested to determine their worth. Rather, the value of specific events is determined according to the shared meanings ascribed to them by actors in the social situation. The same event will, most likely, assume different values (and degrees of worth) in alternate situations. This view of evaluation closely parallels some conceptions of aesthetic criticism in the world of art.

Scriven's (1967) point, that the worth of objectives should be assessed as part of curriculum evaluation, can be accepted within Apple's framework, provided the evaluator takes steps to become cognizant of the shared values and paradigmatic criteria held by the actors within a specified educational (and social) context.

The third idea from general evaluation theory has become common currency in evaluation since Cronbach's



(1963) classic paper in which he urged that evaluation should play a positive role in course improvement. This claim has had the effect of changing the countenance of educational evaluation somewhat, so that a viable model must accommodate the curriculum development process.

From the field of art education (curriculum development and evaluation) several prominent ideas emerged. The first of these is Eisner's (1973) observation that specific behavioural objectives need not be stated in advance to direct meaningful experiences in art. In fact, he suggested that such specification may even be pernicious to the natural course of artistic learning. He proposed alternative types of objectives (expressive and Type III) that do not define too closely the nature of the outcomes.

The force of this argument has served to shift the focus of evaluation in art education. Examination of the congruence of objectives and outcomes becomes subordinate to judgement of the quality of experiences gained. The evaluative focus is reflective rather than pre-ordinate.

The second important idea from art education is that curriculum (or sequence of activities) need not proceed in a pre-determined "logical" fashion to be educationally valuable. It is the quality of the experiences gained, rather than sequence, which determines their value. (Eisner, 1973; Stewart, 1972; Wilson, 1974).





Many art curricula reflect this premise, specifying scope and nature of activities, while determination of particular sequences is left relatively unspecified. Classroom activity is intended to display a purposive rather than purposeful character.

The first of these ideas is concerned with ends (objective) and the second with means (procedure). The two are obviously related and may be considered independently or together in curriculum design. The thrust of these two suggestions is that ends and means should exist in an interactional relationship in art curricula.

In reply to a paper by Popham (1973), Aoki (1973) suggested that different problems of curriculum design were involved in situations where ends and means were not tidily arranged into a "pre-charted" course. He responded to Popham's observation that pre-charting one's course was a time honoured mode of intelligent operation by proposing that the artists' method of deliberately "withholding decision on the goal or objective" with a view to "attending to awareness of the random drift situation" could also be considered a mode of intelligent operation. (p. 5)

Also related to these two observations about means and ends is a third idea relevant to "artistic growth". Beittel (1974), D'Amico (1966), Eisner (1973), Kaelin (1969), Stewart (1972), and many others have reiterated a commonly voiced belief of art educators that identical outcomes



from different students are not necessarily a highly desirable occurrence. Rather, idiosyncratic responses are preferred. Actions, products, and opinions of students that demonstrate growth along individual paths should be sought in art classes. The important implication, for evaluation, of this idea is that "normative" assessment methodologies cannot appropriately be applied to determine "growth".

Salient ideas from the fields of aesthetics, hermeneutics, phenomenology, and anthropology were outlined in some detail in Chapter IV. These related primarily to alternate research methodologies; their purposes, methods and conceptions of validity. Further discussion of methodology is conducted in the second section of this chapter.

#### Pilot Testing of the Model

A prototype of the model (see Appendix "A") was used to conduct an evaluation of the 1975 Fine Arts Council Conference. (The evaluation was requested by the Conference Committee Chairman in October 1974). The benefits anticipated for this study from the evaluation were:

- i) To verify the importance of the model components.
- ii) To test the utility of the model.

As particular conference issues were not specified



by the Conference Committee or Executive of the ATA the foci for the investigation were identified by attending to the issues that emerged at preliminary meetings of the planning committee. Components of the tentative model served to guide efforts to identify issues.

This writer was able to attend all preliminary meetings of the planning committee. Opportunities consequently became available for discussion with the members regarding general issues related to the conference. In addition, the author submitted an agenda item at the last meeting before the conference to elicit conference purposes. Subsequently ten questions were generated which represented the foci of the evaluation. These questions were:

1. What needs did the conference serve?
2. What were the participant's perceptions regarding the conference?
3. What are the potential effects of the conference sessions upon participant's subsequent professional activities?
4. What recommendations for change were made by participants?
5. What were the instructors' perceptions regarding interest generated by specific sessions?
6. What degree of participation was observed at specific sessions?
7. How effectively were the conference plans



implemented?

8. Was the content of specific sessions congruent with participant's expectations?
9. Did the conference achieve its purposes?
10. How valuable was the conference?

Explanation and justification for these questions is supplied in Appendix "B".

Data were collected through interview, questionnaire, document analysis, and observation. Interviews were used to answer specific evaluation questions and identify further issues arising from the conference. The questionnaires were employed primarily to seek participant opinion regarding achievement of conference purposes. As well, general reactions and recommendations for change were sought. The function of observation was to answer some specific questions, regarding reasons for attendance and value of the conference, as well as to identify significant features of the conference not specifiable in advance. Document analysis of registration forms and the Constitution of the Fine Arts Council was conducted to collect data about attendance, rationale and purposes of the conference.

Some difficulties were encountered with respect to data collection due to a shortage of interviewer/observers. As well, a follow-up questionnaire was not distributed due to the pre-Christmas mail strike in 1975.





The final report of the evaluation was submitted to the ATA in February 1976. The conclusions and recommendations are included in Appendix "C".

### Critical Analysis of the Pilot Study

Although it was recognised that the conference did not represent the "usual" situation in which the model is intended to be used, sufficient similarities existed to allow testing of its overall structure. The major differences between a conference "curriculum" and the more "usual" educational programme are length and sequence of activities. For the purposes of the pilot test the conference was regarded as a "mini curriculum".

The two major purposes of the pilot test were to verify the prototype model's components, and to test its utility in a curriculum development situation. Components of the model presented in Appendix "A" were found to define adequately the significant elements of the conference programme, and were consequently useful indicators of conference issues. The useful components of the model were identified as Decision, Rationale, Content, Agents, Facilities (which became "Resources" in the revised model), Implementation, and Outcomes.

The utility of the model, on the other hand, was found to be lacking in several respects. The major weakness uncovered by the pilot test was that the relationships operating between components of the model did not adequately accommodate curriculum development or



evaluative processes. Consequently the utility of the model was limited in terms of its ability to provide an evaluative framework.

Following are the important difficulties encountered that were attributable to the format of the model:

1. The model structure implied linearity of the programme (curriculum) development process. Arrangement of the components suggested a logical development of curricula beginning with determination of rationale and objectives and followed by arrangement of content designed to achieve those objectives. The conference demonstrated very clearly that assumptions of "logical" programme development were not appropriate in this situation. Although the dimensions of the model were found to be present they emerged in ever changing juxtapositions and sequences. The purposes of the conference, for example, were not stated until the last meeting before the event, after the content had been determined. Even then they were defined only at the request of the evaluator.

2. Because the "live" curriculum development situation did not approximate its theoretical counterpart in the model, attempts to conduct an evaluation by determining logical contingency of components and congruence of intentions and outcomes became somewhat procrustean. Instead, attending to issues that emerged "within the components" appeared to be a more pragmatic approach than searching out questions of congruence and logical



contingency.

3. The model was found not to possess an "open framework" as desired. (i.e. independence from a frame of reference designed to pre-determine specific variables for selective analysis.)

Individual components were expressed as discrete entities arranged in a fixed, "idealized" relationship to one another. As with a large number of other current models of evaluation the structure prescribed a model for curriculum development, rather than a model of the process. Consequently, because it represented an "ideal type", it was a closed system.

4. The model did not adequately recognise the social world of the curriculum developers and participants (actors). Procedures of evaluation found to be most useful required examination of the taken-for-granted assumptions and paradigmatic criteria shared by the participants.

On the basis of the above points the model was subsequently modified. Following is the rationale formulated subsequent to the pilot study and examination of the literature. This rationale constituted an important basis for revision of the model.

#### Rationale Directing Model Revision

A survey of the literature, reported in Chapter I, revealed three characteristics thought to be required of an evaluation model appropriate to the visual arts. They



were availability of i) an open ended frame of reference, ii) interpretive techniques, and iii) analysis methods that preserve unique qualities of individual behaviours and events. In addition to these characteristics three other qualities were held to be important for any useful model of evaluation. These were i) Flexibility of viewpoint (appropriate for use "within" the classroom or "external" to it), ii) Flexibility of role (appropriate for formative or summative type assessments), iii) Flexibility of methodology. These qualities were not felt to be satisfactorily demonstrated by the prototype of the model. A discussion of these qualities follows:

Flexibility of Viewpoint. Two specific viewpoints were evident in criticisms levelled, by art educators, at traditional evaluation methodologies. These criticisms were directed at methods of data collection (and analysis) "within classroom" on one hand, and at philosophies of curriculum (and development) "outside the classroom" on the other. A clear distinction should be made at this point between these two aspects of curriculum.

The curriculum development process encompasses a variety of decision making activities that usually occur outside the classroom. These decisions relate to such questions as the nature and arrangement of content, rationale, expected outcomes, facilities and personnel, methods of diffusion of the finished material and so on. Each of these decisions is likely to affect, in some way, the eventual classroom interaction. The product of this activity is most





frequently a set of instructional materials placed in the hands of potential users.

Within the classroom context individual teachers make decisions about appropriateness of specific content to be used, sequencing of material, resources necessary, methods of implementing the programme, acceptability of the rationale, and which school facilities to use. The teacher's decisions will eventually be reflected in specific classroom interactions, student learning and perhaps student products.

Curriculum evaluation, therefore, may be regarded as having a double-barrelled focus represented by curriculum materials and curriculum outcomes. Therefore, the materials (and conditions surrounding their origin) as well as the products of their use in the classroom are subject to evaluative examination. Each of these dimensions is intimately related to the other, making clear separation for the purpose of research difficult and often imprudent.

That the final form of the model would accommodate these twin considerations was regarded as a necessary condition for its viability. The ideal model configuration was held to be one that would define the elements of curriculum evaluation from either perspective ("within" the classroom or "outside" it).

Flexibility of Role. Scriven's (1967) distinction between formative and summative roles of evaluation was also kept in mind. The summative function in curriculum



evaluation requires somewhat different data from formative assessments. A viable model should contain sufficient flexibility to accommodate both roles of evaluation.

Flexibility of Methodology. Another proposition of this study was that a curriculum evaluation model appropriate to the visual arts should logically be based upon epistemological assumptions drawn from an artistic paradigm. This proposal was held as a general principle throughout construction of the model. However, it was recognised that evaluative activity designed to play a formative role during development, dissemination and implementation of a curriculum necessarily demands a broad spectrum of data. Not all of these data are related to the content of the curriculum. In addition, the complex human actions that are determined within what may be defined as the universe of activities related to the curriculum in question are not restricted to examination by one methodology alone. Consequently, other research paradigms are recognised where appropriate, although the interpretive mode remains pre-eminent.

## Part II: Explanation of the Revised Model

The original model was strongly influenced by Stake's matrix. It also reflected the "stages" of curriculum development demonstrated by the Alkin, Provus, Stufflebeam and Taylor/Maguire models. Because the notion



of an "ideal type" of curriculum development proved inappropriate as a basis for directing evaluative action, the revised model was developed with no prescriptions regarding "stages" of evaluative activity. Instead the dimensions of the model describe:

- i) The fundamental evaluative relationship of actors and evaluator in terms of the notion of "Paradigmatic Criteria".
- ii) The elements of curriculum.
- iii) The product of evaluative activity called "Qualitative Exposition".

These dimensions are arranged in a fashion that is analogous to a model of art criticism. Before this analogy is explored however, it is necessary to define the dimensions of the revised model.

#### Dimensions of the Revised Model

Paradigmatic Criteria. These criteria are employed by the actors and evaluator for different (but related) ends. They are used by actors to choose between curriculum alternatives related to questions such as procedures, content, resources, etc. The evaluator uses his paradigmatic criteria to select foci for evaluative attention and to judge the value of observed educational phenomena.

These criteria have their origins in the "personal knowledge" (Polanyi, 1964) of individuals who are



concerned with making judgements in the educational context. They reflect the individual's "world view", which is determined by his biographical situation (Schutz, 1973). The expressed form of these criteria will frequently constitute shared meanings and values, held by actors within the social (curriculum) group. The degree of congruence that exists between the two sets of criteria possessed by actors and evaluators will vary greatly, due to such determining background factors as cultural origin, philosophic orientation, educational background and so on.

In truth, many more than two sets of paradigmatic criteria will probably exist. Amongst the "actors" differences among criteria employed by curriculum developers and various sub-groups of teachers will be unavoidable. Schutz' notion of "multiple realities" is a useful explanatory tool for comprehension of this occurrence.

The frame of reference for the evaluator and actors in the educational context is the curriculum, or curriculum development process. Curriculum has been defined in terms of six irreducible elements. Each of those elements is described below from the two perspectives outlined previously in the discussion of the rationale for the model; i.e. "external" to the classroom (the perspective of the developers) and "within" the classroom (the perspective of the classroom participants).

Curriculum Elements. These are discussed in





alphabetical order:

- i) Agents refers to the personnel of the programme. "Agents" within the classroom are the teachers or "initiators of experiences". External to the classroom agents are the programme developers and administrators. "Developers" include individuals representing government or social agencies and educational organizations. Evaluative questions which could be raised with respect to agents include issues like qualifications of curriculum developers, vested interests of specific individuals, teacher preparation necessary for implementation of programmes etc.
- ii) Content is the subject matter of the programme including such considerations as objectives (general and specific), activities and topics. Outside the classroom the "content" means curriculum materials. Within the classroom however, content will vary according to the teachers' individual choices, and interpretations of written curriculum materials. Possible evaluative issues could be related to suitability, nature, clarity or sophistication of this content.
- iii) Implementation carries dual meanings depending upon the perspective from which it is considered. From the viewpoint of curriculum



development, "implementation" means disseminating of curriculum materials (or philosophies). From the classroom perspective implementation is the translation of curriculum materials into instruction. Evaluative issues connected with implementation could be appropriateness or effectiveness of distribution methods, and instructional "style".

iv) Outcomes are results of the programme manifested as transactions and products. From within the classroom cognitive, affective and psychomotor dimensions of experience are probable sources of evaluative examination. Outcomes are not limited to students; teachers may demonstrate programme effects as well. From a position external to the classroom an outcome of the curriculum development process could be regarded variously as a unit of written curriculum materials or statements of "expected" outcomes. In all respects the outcomes of the programme are the most significant indicators of its success. The evaluative issue is the omnibus question concerning the ultimate value of the programme. All other evaluation questions are directed towards understanding the reasons for the outcomes being as they are.

v) Rationale is the explanation of reasons



for following a certain course of action. It is recognised that the rationale may be implicit or explicit. However in this context "Rationale" refers to explicit statements of reasons for action. Outside the classroom, at the curriculum development level, political, societal, and current trends in educational philosophy are likely to be more readily reflected in the rationale than at the classroom level. Within the classroom, teachers' personal knowledge of curriculum content and the immediate needs of students will be likely to result in variations from "official statements" of curriculum "rationale". Evaluative issues connected with rationale will reflect questions of relevance or "appropriateness".

vi) Resources are the necessary accompaniments of curriculum. Outside the classroom funding heads the list followed by items such as facilities, instructional aids, and specialist personnel. Within the classroom the resources question is usually related to administration (i.e. How to get specific items?). The usual evaluation question about this element is cost/benefit related.

Qualitative Exposition. This is the product of the evaluative process. It is a verbal, and/or visual presentation of the evaluator's findings, possessing characteristics which are similar to those of "educational criticism" outlined by Eisner (1975). They are:

i) Use of vividly descriptive techniques



(language and/or visual).

ii) Demonstration of evidence of the evaluator's paradigmatic criteria through analytical, interpretive and judgemental statements.

The differences between scientific and artistic language are profound. Whereas quantitative description contains nothing which might suggest an inherent structural relationship between the object being described and the language used to describe it, artistic language actively seeks to employ such relationships. Artistic language is so phrased as to convey a certain quality of life which may be said to permeate the experience being described. Its impact is visceral, and its form is expressive of the experience itself.

#### Relationships of Model Components

In the statement of Purposes of the Study it was hypothesised that an appropriate framework for evaluation of an art curriculum should properly be derived from the artistic paradigm. The process of art criticism was therefore examined in terms of its potential for translation into a method of curriculum evaluation.

The following parallels were noted between the two activities.

i) The roles of "art critic" and the "curriculum evaluator" are essentially the same. Both individuals collect data about an





object or event and interpret their meaning for a specified audience.

ii) Both curricula and art objects depend for their structure and content upon hundreds of decisions made by their creators.

iii) The forms of both phenomena are characterized by a discrete number of generic elements.

iv) Both "art" and "curriculum" are open concepts. That is, there is no set of necessary and sufficient conditions which can adequately define "good" and "bad" art or "good" and "bad" curricula.

v) Both art criticism and curriculum evaluation are valuative activities. The art critic draws upon his internalized knowledge of artistic phenomena to make judgements about the worth of a work of art. The curriculum evaluator draws upon his widely differentiated knowledge of educational theory and practice to assess the value of a programme.

The major difference between art and curriculum was discussed above. Art objects are inanimate and do not change their form. Curricula, on the other hand, are dynamic social constructions.

A Proposed Model for Art Criticism. To develop further the analogy between art criticism and curriculum evaluation, a model for art criticism (for two-dimensional



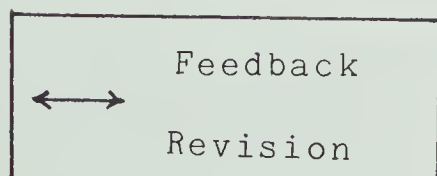
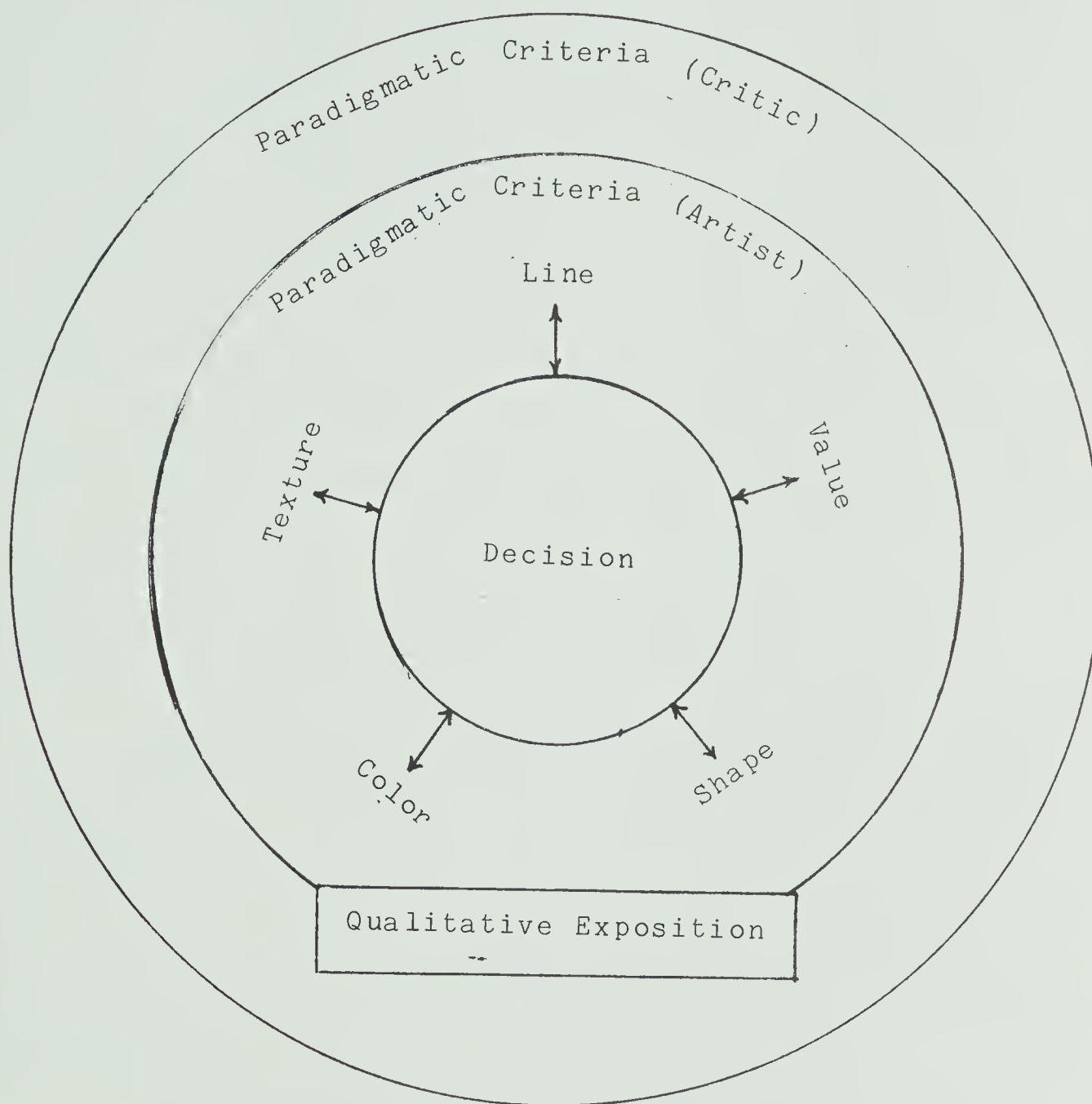
works) was proposed. (See Figure 2). The simple premise upon which this model was constructed is; that art criticism requires attention (by the critic) to the art work for the purpose of making judgements of its value. These judgements and the reasons that support them are clearly communicated to an audience. (Beardsley, 1971; Dewey, 1934; Ecker and Kaelin, 1972; Suniga, 1973). In addition to this premise the following assumptions were held to be sound:

- i) That a two-dimensional art work is produced by manipulation of all or some of the elements of line, shape, value, texture and colour.
- ii) That decisions about specific manipulations, by the artist, of these elements determines the final form of the object.
- iii) That decision making is central to the production of an art work.
- iv) That the artist's "paradigmatic criteria" will be embodied in decisions about specific manipulations of the elements.
- v) That the critical reasons (Beardsley) offered by the critic in support of his judgement of an art work will reflect his "paradigmatic criteria".

The form of the model presented in Figure 2 demonstrates the process of art criticism as the juxtaposition of two sets of paradigmatic criteria;



Figure 2.  
Model for Art Criticism  
(Two Dimensional Works)





the critic's (represented by the outer circle) and the artist's (represented by the second circle). The subject of criticism (the art work) is represented in terms of its minimal elements fashioned, through decision making processes (the centre circle) into a unique relationship. The critic, through prior knowledge, will be aware that the artist attended to each element, either separately or together, for the purpose of gaining feedback. Decisions made, on the basis of that feedback, resulted in the revision to the arrangement of elements until the product was regarded by the artist as complete.

The methodology of art criticism requires an act of refined perception (Eisner, 1975; Feldman, 1970) which necessarily implies skills of interpretation and analysis. (The epistemology of art was discussed in Chapter IV in connection with concepts of validity).

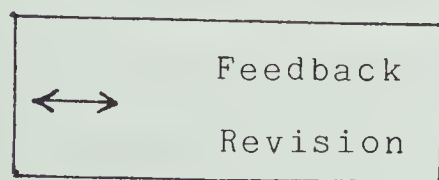
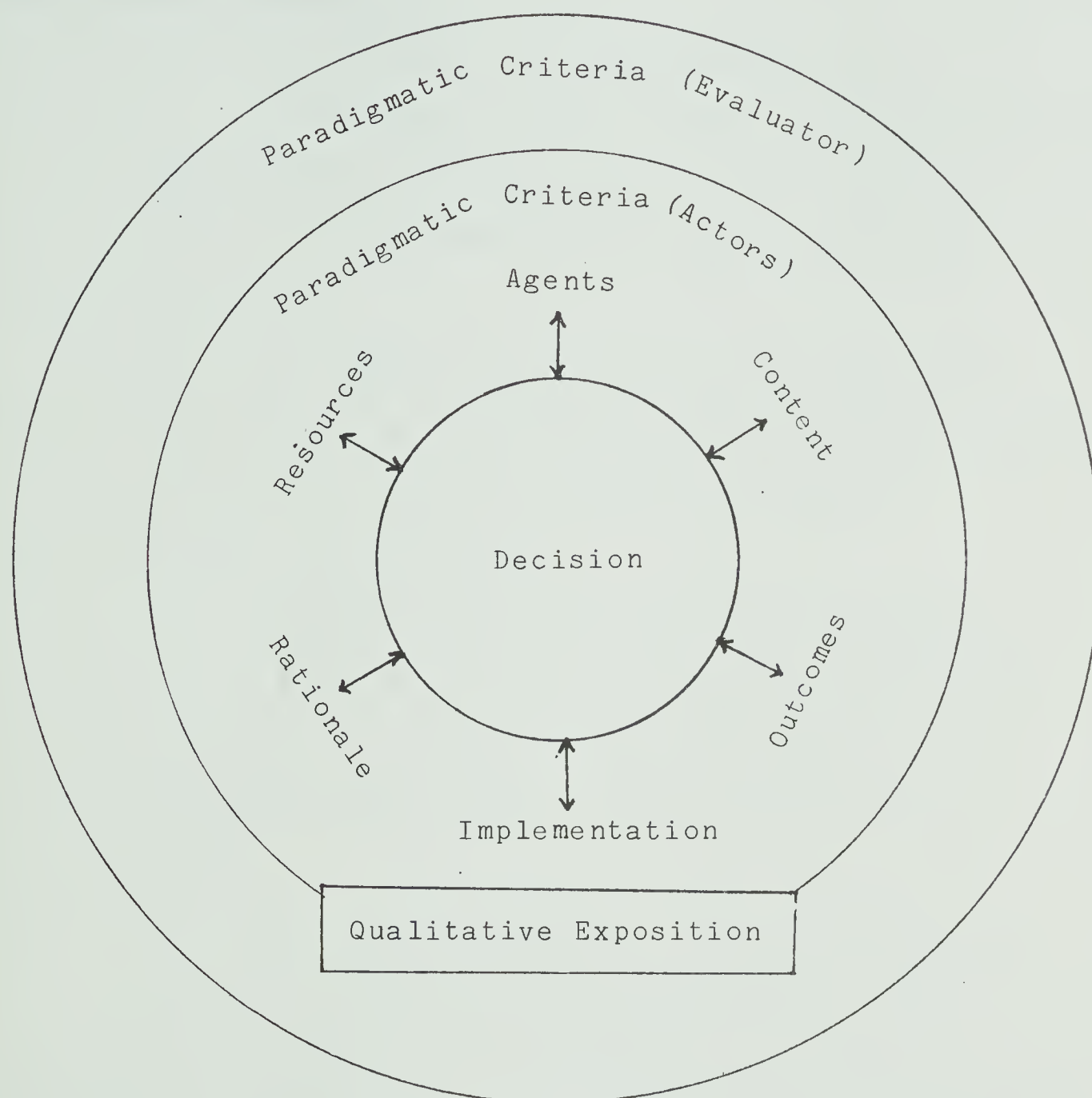
The Revised Curriculum Evaluation Model. Because of the similarities between art criticism and curriculum evaluation that were noted before, the framework of the model of art criticism was used as a pattern to formulate the revised model for curriculum evaluation in the visual arts (see Figure 3).

The structure of the model remained the same while the following terminological substitutions were made. Within the Paradigmatic Criteria dimensions "critic" and "artist" were replaced by "evaluator" and "actors" (curriculum developers, teachers, and students). The





Figure 3.  
Curriculum Evaluation Model  
for the Visual Arts.





elements of two-dimensional art work were replaced by the elements of curriculum described before as Agents, Content, Implementation, Outcomes, Rationale, and Resources. The central Decision component remained, as did "Qualitative Exposition".

The implications of this arrangement are:

- i) Evaluation includes the act of juxtaposing the paradigmatic criteria, held by the evaluator, against those of the actors in the curriculum situation.
- ii) The object of evaluation (the curriculum) is a product of the actor's decision-making activities relative to any or all of six elements of curriculum.
- iii) Decisions made about curriculum elements will largely determine the ultimate form of the curriculum. As well, these decisions will embody the paradigmatic criteria held by the actors.
- iv) Curriculum elements are not specified for evaluative attention in any particular order.

By extending the analogy of art criticism to curriculum evaluation the "open framework" characteristic of the epistemology of art is preserved. This is evident in two ways. First, it is proposed that features of curriculum are selected, examined, and judged in light of the evaluator's paradigmatic criteria. Mann's (1969)



observation regarding the role of personal knowledge in the initiation of research is usefully adapted for clarification of this point: A research project begins with a decision about what to select for attention, and that decision is governed by what one's personal knowledge leads him to believe will be valuable and fruitful. The evaluator's problem is to select, from an inexhaustible realm of designs and meanings, those he will attend to, and that decision is grounded in personal knowledge. The evaluator approaches the phenomena to be examined with a set of predispositions, in the form of highly abstract models of what designs it would be of value to discover. These "models" can be regarded as his "paradigmatic criteria".

Second, although the elements of curriculum are defined by the model, they are not specified to exist in any pre-determined relationship. The nature of their relationship is determined by application of the paradigmatic criteria shared by the "actors". Because the model does not delineate an "ideal type" for curriculum or curriculum development the evaluative questions are not pre-determined.

The major translation from a model of art criticism to curriculum evaluation is necessary in the area of observation methodologies. As was mentioned before, the structure of the proposed model is analogous to the framework of aesthetic criticism, and the primary mode of



inquiry in that context is interpretation of meaning. The obvious difference between art criticism and curriculum "criticism" lies in the nature of the object under investigation. The art object is inanimate; it attributes no meaning to its own structure, whereas a "curriculum" is comprised of an ever changing system of relationships among people, objects and events. To complicate the issue, people manufacture their own meaning structures to explain their relationships to the system.

To derive data from an art object requires a simple process of observation; the object pays the viewer the courtesy of holding still (usually). Observation of the social construction that represents "curriculum" is a much more complex task. A discussion of interpretive methodologies in the social situation follows in the next section.

### Methodological Assumptions

Evaluative methodology within this model is defined by the relationship of "evaluator" and "actors". In the previous discussion of the function of the evaluator's paradigmatic criteria it was noted that the evaluator selects for observation, and judges, aspects of curricula in terms of those criteria.

As the evaluator attends to the curricula he must attempt to gather a wide variety of data for interpretation.





The specific data gathered will be determined by his initial decision about which aspects of curricula to attend to. The primary mode of his investigations will be interpretive (valuative) in order that he may achieve the desired evaluative product - a "qualitative exposition".

Although it is recognised that quantitative data may need to be collected in connection with each of the elements of curriculum, the most significant implication of the "paradigmatic criteria" component of the model is evident in examination of the "outcomes" in the classroom context.

Because this model of evaluation does not demand pre-specification of particular objectives in art education no difficulty is presented in the event that curriculum developers or teachers choose not to do so. The task of the evaluator is to determine the value of experiences gained by the students. This requires exercise of judgement with respect to selection of outcomes considered to be significant and determination of their worth. The "paradigmatic criteria" of the evaluator are employed to this end.

The obvious critical issue that is raised in connection with this type of methodology is that of validity of judgements made. How can one be sure that the issues and judgements are as the evaluator says they are? This question is discussed in the next section.

The kinds of outcome data that may be collected



will again vary according to the evaluator's interests. They fall roughly into two classes, which may be called "objective" and "subjective" data. Objective data are external information about curriculum outcomes and will include items such as video-tapes of lessons, audiotapes, observer notes and student products. Subjective information will consist of actors' explanations of their reasons for actions and classroom critics' judgements of the value of observed behaviours. Both types of data may be used to support evaluative judgements. Writers such as Eisner (1975), Pohland (1972), Stake (1975) and Schatzman and Strauss (1973) have recommended collection of the widest possible data base upon which to form the most useful judgements. As well, variety of data (visual and verbal) will add to the vividness of the report to decision makers (Eisner, 1975).

Data may be analyzed in two distinct ways; in order to discover causes of actions or to discover "reasons". This distinction was discussed in Chapter IV. It is argued that analysis in terms of "reasons" in the phenomenological mode is the approach most compatible with the epistemology of art. Such analysis accepts free will of the individual, which is a basic tenet of artistic behaviour. However, both kinds of analysis are useful for different purposes.

Other kinds of data (and analysis) will be appropriate for different elements of curriculum. To determine an appropriateness of content in a set of



curriculum materials, for example, a panel of "expert judges" could be consulted. The empirical - analytical (consensus) mode would be the logical choice in this instance. Quantitative data about resources, implementation procedures and agents are also likely to be required to answer common evaluation questions (like "How many teachers have art backgrounds?" or "How much visual material is available to the school?").

The central methodological issue has not been developed fully here. In the next chapter the methods employed to collect and analyze outcome data are explained in detail within the context of the Edmonton Public School Board Art Appreciation Project. In this project specific behavioural objectives were not specified by the curriculum developers, so necessitating the use of techniques of reflective analysis and judgement of the value of outcomes. The reasons supplied to support the decisions made in this evaluation will serve to demonstrate, in a practical way, the methodology implied by the model.

### Validity of Implied Research Methodologies

There are three essential conceptions of validity embodied within the methodology of evaluation in this model. Each of these relates to the kind of data used to draw conclusions. These possible data types are empirical, judgemental, and phenomenological.

Empirical data have been commonly used in educational



and evaluative research and require little further elaboration. Validity controls rely upon consensus techniques.

Judgemental data are produced most commonly in the context of classroom observation. In the usual art lesson desired outcomes are not specified explicitly in advance; therefore the judgemental process requires the evaluator to i) select those aspects of the transactions or products he considers to be important and ii) assess the worth of these outcomes. This is an analogous process to art criticism, as discussed in Chapter IV. The reader will recall that the criterion considered most useful for validity of judgement in that context was "referential adequacy". That is, judgements made must be supported by evidence supplied from the object under investigation. Eisner (1975) has used this concept in the context of his discussion of "educational criticism" and "educational connoisseurship".

The reader will also recall that reliability does not constitute a criterion of validity in art criticism as it does in empirical-analytical research. It is possible for different individuals to attend to alternative aspects of the same phenomenon and support their judgements equally with various evidence. Mann attributes differences in critical judgement to the fact that the critic is required to draw upon his personal knowledge (of relevant appropriate or worthwhile designs) in order





to select which of an infinite range of possible patterns to attend to. Harris (1973) attributes the difference to the relativity of perception and differing educational backgrounds of observers. Ecker and Kaelin claim that the responsibility of the audience is to check their own perceptions with those of the critic.

Child (1906), on the other hand, has conducted considerable research in the area of aesthetic preferences, and has discovered that experts tend to agree far more, in matters of aesthetic judgement, than do laymen. It is not unreasonable to argue then that in art education, experts would be likely to agree also. Nevertheless variety of viewpoint is an advantage, rather than a disadvantage to the evaluator.

Harris suggested that aestheticians should revel in the unique and varied ways that people discuss their art experiences; they should emphasise the uniqueness of each description. This argument could also be extended to curriculum evaluation. The value of several sets of judgements to the evaluator lies in the richness of the data base provided.

The preceding discussion of validity of judgemental data was predicated upon the assumption of "detached" observation. The judge's information is based upon external (non-participant) information. If the classroom observer becomes involved with the situation, in terms of questioning participants about their reasons for action, an entirely



different set of criteria for validity becomes effective.

To explain human actions in terms of the actor's own reasons is a phenomenological approach. The basic criterion for validity in this context is that the actor understand, and agree with the researcher's explanation of the behaviour (The Postulate of Adequacy). This concept of validity was also explained at length in Chapter IV.

Each of the three above methods provide particular, useful evaluative information. Each is likely to be used to some extent in almost every evaluation project because evaluations are essentially complex endeavours.

The overriding question is that of the evaluator's choice of focus. Although the evaluator may support his conclusion with evidence of proven validity his choice of questions must also be responsible.

If the evaluator's questions are not viewed by the decision-makers as useful or meaningful then the findings will be irrelevant. Therefore a "valid" evaluation question has to be defined as that which carries intersubjectively shared meaning for the evaluator and the decision-makers. This will require the evaluator to discover the shared values of the actors. The source of this idea is derived from Apple's suggestion, mentioned previously.

### Summary of Chapter V

In the first section of this chapter the steps



taken to develop a curriculum evaluation model for the visual arts were described. They were: i) analysis of the components of curriculum development in two live projects; ii) selection of salient ideas from the literature; iii) pilot testing of the first form of the model; iv) revision of the model in light of the pilot test.

The rationale which guided revision of the model was outlined. Features of a viable model (in addition to those stated in the Purposes of the Study) were held to be: i) flexibility of viewpoint; ii) flexibility of role; iii) flexibility of methodology.

Data sources, both theoretical and empirical, were described. Salient ideas from the literature were discussed. From general evaluation theory three central ideas were identified. These were: i) evaluation is a process of defining, obtaining, and providing useful information to decision-makers; ii) evaluation is a process of social valuing; iii) evaluation should play a role in curriculum improvement. From art education and evaluation of art three further ideas were isolated: i) in art programmes specific objectives need not be stated in advance to ensure qualitative programmes; ii) lesson content in art programmes does not necessarily have to proceed in a pre-determined fashion to be worthwhile; iii) desired student responses in art are individual rather than "normative".



Pilot testing of the model revealed the prototype of the model was successful in identifying evaluative issues related to model components, but was unsuccessful in terms of encapsulating curriculum development and evaluative processes. The reasons identified for the poor utility of the model were i) linearity; ii) inappropriateness of the evaluative question of logical contingency and congruence; iii) absence of an open framework; iv) lack of recognition of the social dimension of the evaluative process.

In the second section of this chapter the revised model was explained. First, the dimensions of the model were delineated. They were i) Paradigmatic Criteria; ii) Curriculum Elements (agents, content, implementation, outcomes, rationale, resources); iii) Qualitative Exposition.

Second, the relationships between the components of the model were considered as analogous to the structure of the process of art criticism. This analogy was explored at length. The central proposition of this discussion was that construction of a model of curriculum evaluation that was analogous to the method of art criticism would preserve an open frame of reference.

Methodological implications of the model were briefly discussed. Two types of outcome data were identified as objective and subjective. Analysis of data was observed to be possible in terms of causes or reasons. It was





argued that the mode of research most compatible with the structure of the model was phenomenological. Use of other modes, however, was also discussed.

Three essential conceptions of validity were outlined in the final section of the chapter. They were i) consensus; ii) referential adequacy; and iii) the phenomenological notion of the actor's recognition of explanations (Postulate of Adequacy). These three conceptions respectively are identified with empirical-analytical, artistic, and phenomenological research methods.



# REFERENCES

- Aoki, T. A Response to Dr. W. James Popham's Paper entitled "Curriculum Design: The Problem of Specifying Intended Learning Outcomes". Presented at a symposium on program development in education at the University of British Columbia, May 24, 1973.
- Apple, M.W. The Process and Ideology of Valuing in Educational Settings. In M.W. Apple, M.J. Subkoviak & H.S. Lufler Jr., (Eds.), Educational Evaluation: Analysis and Responsibility. Berkely: McCutchan, 1974.
- Beardsley, M.C. The Classification of Critical Reasons. In R.A. Smith (Ed.), Aesthetics and Problems of Education. Chicago: University of Illinois Press, 1971.
- Beittel, K.R. Formative Hermeneutics in the Arting Processes of an Other: The Philetics of Art Education, Art Education, 1974, 27(9), 2-7.
- Child, I.L. The Problem of Objectivity in Esthetic Value. Penn State Papers in Art Education, 1966, 1.
- Cronbach, L.J. Evaluation for Course Improvement. Teachers College Record, 1963, 64, 676-683.
- D'Amico, V. Millennium or Mirage. Art Education, 1966, 19(5), 27-32.
- Dewey, J. Art as Experience. New York: G.P. Putnam's Sons, 1934.
- Ecker, D.W., & Kaelin, E.F. The Limits of Aesthetic Inquiry: A Guide to Educational Research. In P.L.G. Thomas (Ed.), Philosophical Redirection of Educational Research. The Seventy-First Yearbook of the National Society for the Study of Education. Part I. Chicago, Illinois: University of Chicago Press, 1972.
- Eisner, E.W. Do Behavioural Objectives and Accountability Have a Place in Art Education? Art Education, 1973, 26(5), 2-5.
- Eisner, E.W. The Perceptive Eye: Toward the Reformation of Educational Evaluation. (Unpublished mimeographed paper of an address presented at a seminar, University of Alberta).



- Feldman, E.B. Becoming Human through Art. Englewood Cliffs, New Jersey: Prentice Hall, Inc., 1970.
- Harris, I.M. Agreement in Art: An Educational Issue: Journal of Aesthetic Education, 7(3), 1973, 51-61.
- Kaelin, E.F. Are Behavioural Objectives Consistent with Social Goals of Aesthetic Education. Art Education, 1969, 22(8), 4-11.
- Mann, J.S. Curriculum Criticism. Teachers College Record, 1969, 71(1), 27-40.
- Pohland, P. Participant Observation as a Research Methodology. Studies in Art Education. 1972, 13(3).
- Popham, W.J. Curriculum Design: The Problem of Specifying Intended Learning Outcomes. Presented at a Symposium on program development in education at the University of British Columbia, May 24, 1973.
- Schatzman, L., & Strauss, A.L. Field Research: Strategies for a Natural Sociology. Englewood Cliffs, New Jersey: Prentice Hall, Inc., 1973.
- Schutz, A. Common Sense and Scientific Interpretation of Human Action. In M. Natanson (Ed.), Alfred Schutz: Collected Papers I: The Problem of Social Reality. The Hague: Martinus Nijhoff, 1973.
- Scriven, M. The Methodology of Evaluation. In R.E. Stake (Ed.), Curriculum Evaluation. American Educational Research Association Monograph Series on Evaluation, 1, Chicago: Rand McNally, 1967.
- Stake, R.E. The Countenance of Educational Evaluation. Teachers College Record, 1968, 68, 523-540.
- Stake, R.E. To Evaluate an Arts Program. In R.E. Stake (Ed.), Evaluating the Arts in Education: A Responsive Approach. Columbus, Ohio: Charles E. Merrill Publishing Company, 1975, 13-32.
- Stewart, W. A Humanistic Study of the Visual Arts, The Commitment, Risk and Potential. Art Education, 1972, 24(2), 2-3.
- Stufflebeam, Foley, Gephart, Guba, Hammond, Merriman, & Provus. Educational Evaluation and Decision Making. Itasca, Illinois: F.E. Peacock Publishers, Inc., 1971.



- Wilson, B.G. The Other Side of Evaluation of Art Education. In G.W. Hardiman & T. Zernich (Eds.), Curricular Considerations for Visual Arts Education: Rationale Development and Evaluation. Champaign, Illinois: Stipes Publishing Company, 1974, 247-76.
- Worthen, B.R., & Sanders, J.R. Educational Evaluation: Theory and Practice. Worthington, Ohio: Charles A. Jones Publishing Company, 1973.
- Zuniga, J.A. Critical Analysis and its Subject Matter. Journal of Aesthetics and Art Criticism, 1973, 32(2).





## CHAPTER VI

### Validation of the Model

The primary purpose of this chapter is to demonstrate the validity of the model. Both empirical and theoretical evidence is employed to this end. An examination of the meaning of validity, undertaken in Chapter IV, demonstrated the evanescent nature of the concept. In general terms, however, validity can be regarded to mean strength or "scientific utility" (Kaplan, 1964). The following explanations and illustrations were considered to contribute towards establishing validity of the model in pragmatic and theoretical terms, if one accepts the three norms for theory validation identified by Kaplan (the norms of correspondence, coherence, and pragmatism).

Major questions raised by these norms are addressed separately. The largest portion of this chapter is devoted to an examination of the results obtained through using the model as a framework to conduct an evaluation of the Edmonton Public School Board Art Appreciation Project. This discussion addresses the questions raised by the norm of Pragmatism. In addition to demonstrating the practical value of the model, a significant benefit of this discussion lies in its



clarification of the methodological procedures that were outlined in general terms in the previous chapter. As well, some of the practical problems relating to use of this methodology are raised.

### The Norm of Correspondence

The central question raised by this norm is: Does the theory fit the facts? Any theory of curriculum evaluation is, in fact, a concatenation of theories, basic to which are notions of "curriculum" and "evaluation". Each of these is a socially constructed concept which is empirically observable only in terms of its operational definition.

As the model of this study represents a proposition for evaluation, rather than a universal explanation of the process there seems no point in attempting to validate it on empirical grounds (i.e. what usually happens).

The model does not, however, exclude descriptive elements which represent components of the process of curriculum development. In other words, contained within the model for evaluation is a sub-unit which comprises a model of the curriculum development process. This sub-model was empirically verified. (See Table 1)

The procedure employed to identify the operational components of curriculum development was outlined in Chapter V. In the course of observation of the early



meetings of the Fine Arts Council Conference Planning Committee and the Edmonton Public School Board Art Appreciation Committee the researcher was able to verify six categories of discussion. These categories were identified in Chapter V as the elements of curriculum development. Table 1 is a summary of selected samples of topics discussed by each committee. These are arranged under the categories identified as Agents, Content, Implementation, Outcomes, Resources, and Rationale. This table serves to indicate the nature of the data which contributed towards verification of the curriculum elements.

A panel of six judges was employed to sort the topics represented in Table 1 into the most appropriate categories. Five of the six judges were members of one or both curriculum committees. The sixth judge was a doctoral student in the area of curriculum. Each independently classified the discussion topics. Although the option to discard any topic was available it was not used.

Judge agreement was determined by calculating separately the average percentage of agreement between judges within each category and for both projects. (See Tables 2 - 13). Table 14 presents a summary of the results. Overall agreement was high, indicating that the categories did, in fact, possess a degree of face validity. Agreement was higher for the EPSB project than the FAC, suggesting that the category definitions may be more appropriate for



TABLE 1

Categorization of Topics Discussed  
by Curriculum Committees

	FAC Conference Committee	EPSCB Curriculum Committee
Agents	<ol style="list-style-type: none"> <li>1. Identification of possible speakers.</li> <li>2. Selection of chairpersons.</li> <li>3. Selection of appropriate personnel for committee responsibilities.</li> </ol>	<ol style="list-style-type: none"> <li>1. Necessary qualifications (in art) of elementary teachers to teach the proposed programme.</li> <li>2. Selection of pilot teachers.</li> </ol>
Content	<ol style="list-style-type: none"> <li>1. Discussion of major content areas to be covered by speakers at individual sessions.</li> <li>2. Selection of conference entertainment.</li> </ol>	<ol style="list-style-type: none"> <li>1. Selection of appropriate objectives, foci and activities for materials.</li> <li>2. Organization of materials.</li> <li>3. Generation of curriculum activities.</li> <li>4. Design of integrative activity.</li> <li>5. Generation of resource references.</li> </ol>
Implementation	<ol style="list-style-type: none"> <li>1. Acquisition of venue.</li> <li>2. Allocation of space.</li> <li>3. Organization of programme.</li> <li>4. Organization of registration procedure.</li> </ol>	<ol style="list-style-type: none"> <li>1. Distribution of material.</li> <li>2. Nature of required in-service.</li> <li>3. Alternative methodologies for use of materials.</li> </ol>





(Table 1 continued.)

	FAC Conference Committee	EPSB Curriculum Committee
Outcomes	<ol style="list-style-type: none"><li>1. Most effective methods to ensure full participant involvement.</li></ol>	<ol style="list-style-type: none"><li>1. Anticipated teacher use of materials.</li><li>2. Anticipated response of students to materials.</li><li>3. Observed response of students to selected materials.</li><li>4. Observed usefulness of materials in an integrated programme.</li></ol>
Resources	<ol style="list-style-type: none"><li>1. Allocation of available budget.</li><li>2. Advantages and disadvantages of available space.</li><li>3. Acquisition of Audio Visual requirements.</li></ol>	<ol style="list-style-type: none"><li>1. Inefficiency of EPSB Instructional Materials Centre.</li><li>2. In-service for the construction of teacher made visual aids.</li><li>3. Funds available for visual aids.</li></ol>
Rationale	<ol style="list-style-type: none"><li>1. Reasons for placing emphasis on workshop sessions.</li><li>2. Possibility of integrating Art, Music, Drama.</li></ol>	<ol style="list-style-type: none"><li>1. Function of materials<ol style="list-style-type: none"><li>a) To serve as mini-course for teachers.</li><li>b) To supplement existing curriculum guide.</li><li>c) To demonstrate "integrative" aspects of art appreciation.</li></ol></li><li>2. Non product orientation of materials.</li></ol>



TABLE 2  
JUDGE AGREEMENT  
EPSB CATEGORY I: AGENTS

	I	II	III	IV	V	VI
I		100	100	100	100	100
II			100	100	100	100
III				100	100	100
IV					100	100
V						100
VI						

$\bar{X}\%$  AGREEMENT = 100%

TABLE 3  
JUDGE AGREEMENT  
EPSB CATEGORY II: CONTENT

	I	II	III	IV	V	VI
I		60	67	67	67	67
II			60	60	60	60
III				100	100	100
IV					100	100
V						100
VI						

$\bar{X}\%$  AGREEMENT = 77%



TABLE 4  
JUDGE AGREEMENT  
EPSB CATEGORY III: IMPLEMENTATION

	I	II	III	IV	V	VI
I		40	50	50	50	50
II			50	50	50	50
III				100	100	100
IV					100	100
V						100
VI						

$\bar{X}\%$  AGREEMENT = 69%

TABLE 5  
JUDGE AGREEMENT  
EPSB CATEGORY IV: OUTCOMES

	I	II	III	IV	V	VI
I		40	100	100	100	100
II			40	40	40	40
III				100	100	100
IV					100	100
V						100
VI						

$\bar{X}\%$  AGREEMENT = 80%



TABLE 6  
JUDGE AGREEMENT  
EPSB CATEGORY V: RATIONALE

	I	II	III	IV	V	VI
I		67	100	100	100	100
II			67	100	100	100
III				100	100	100
IV					100	100
V						100
VI						

$\bar{X}\%$  AGREEMENT = 95%

TABLE 7  
JUDGE AGREEMENT  
EPSB CATEGORY VI: RESOURCES

	I	II	III	IV	V	VI
I		75	50	50	50	50
II			67	67	67	67
III				100	100	100
IV					100	100
V						100
VI						

$\bar{X}\%$  AGREEMENT = 76%





TABLE 8  
JUDGE AGREEMENT  
FAC CATEGORY I: AGENTS

	I	II	III	IV	V	VI
I		75	100	100	100	100
II			75	75	75	75
III				100	100	100
IV					100	100
V						100
VI						

$\bar{X}\%$  AGREEMENT = 91%

TABLE 9  
JUDGE AGREEMENT  
FAC CATEGORY II: CONTENT

	I	II	III	IV	V	VI
I		25	0	0	0	20
II			0	0	0	20
III				100	100	100
IV					100	50
V						50
VI						

$\bar{X}\%$  AGREEMENT = 37%



TABLE 10  
JUDGE AGREEMENT  
FAC CATEGORY III: IMPLEMENTATION

	I	II	III	IV	V	VI
I		25	50	60	60	60
II			33	40	40	40
III				83	83	57
IV					100	67
V						67
VI						

$\bar{X}\%$  AGREEMENT = 57%

TABLE 11  
JUDGE AGREEMENT  
FAC CATEGORY IV: OUTCOMES

	I	II	III	IV	V	VI
I		100	100	100	100	100
II			100	100	100	100
III				100	100	100
IV					100	100
V						100
VI						

$\bar{X}\%$  AGREEMENT = 100%



TABLE 12  
JUDGE AGREEMENT  
FAC CATEGORY V: RATIONALE

	I	II	III	IV	V	VI
I		50	50	50	33	0
II			100	100	67	50
III				100	75	50
IV					67	50
V						33
VI						

$\bar{X}\%$  AGREEMENT = 58%

TABLE 13  
JUDGE AGREEMENT  
FAC CATEGORY VI: RESOURCES

	I	II	III	IV	V	VI
I		67	67	100	67	100
II			33	67	33	67
III				67	100	67
IV					67	100
V						67
VI						

$\bar{X}\%$  AGREEMENT = 71%



TABLE 14  
JUDGE AGREEMENT: SUMMARY TABLE

	AGENTS	CONTENT	IMPLEMENTATION	OUTCOMES	RATIONALE	RESOURCES	$\bar{X}\%$
FAC	92%	38%	58%	100%	58%	71%	69%
EPsB	100%	78%	69%	80%	96%	76%	83%
$\bar{X}\%$	96%	58%	64%	90%	77%	74%	76%





traditional school curriculum development projects than for conference type curricula. Some difficulty was noted with "content" in the FAC project, suggesting confusion with respect to viable topics of conference content.

In addition to the content of discussions the researcher also identified three determining processes that were observed to operate at these meetings. They were feedback, decision, and revision.

Feedback was the process of collecting and supplying data about the curriculum elements to decision-makers. For example, in the winter term of 1975, several of the EPSB Committee members tested some first draft materials on their own classes with a view to collecting feedback information for the rest of the committee. This information served as the basis for generating decision alternatives.

Decision making was the central process intention of all discussions. Many decisions were made at each meeting of both committees.

Revision was the third observed process in curriculum development. Following the decision making process, revisions to curriculum elements often occurred.

In summary, then, the model emerges as an eclectic configuration of empirical and postulational components. Within the total framework, however, the "curriculum development" dimensions were demonstrated to be verified empirically. In response to the question of "fitting



the facts", it can be argued that at least the sub-model possesses a certain degree of empirical validity and therefore fits the norm of correspondence.

### The Norm of Coherence

There are three questions to be dealt with separately here. The first is: How does the model integrate with existing theories of curriculum evaluation?

Before attempting to attend to this question it must be noted that the body of existing evaluation theory should be recognized as a collection of loosely associated sets of hypothetical assumptions which demonstrate varying degrees of overlap with each other. The term "theory" in the context of curriculum evaluation cannot be afforded its strict scientific meaning of "explanation of phenomena". Rather it should more appropriately be regarded in a general "propositional" sense. Most theories (or models) of evaluation represent propositions for the conduct of evaluative enquiry based upon certain assumptions. The curriculum evaluation model for the visual arts proposed in this study is built upon assumptions that are congruent with basic propositions of existing theories.

Significant ideas, derived from the literature, were used as basic principles to develop the revised model. These notions were discussed in the previous chapter under the notation of Salient Ideas.

The first of these ideas, connected with a



definition of evaluation, is particularly significant. The evaluative process was held to consist of delineating, obtaining, and providing useful information to decision-makers, to assist in judging decision alternatives. Three essential components of this process were delineating, obtaining, and providing. Most existing models of evaluation describe or imply these components with greater or lesser degrees of clarity.

The model proposed in this study also observes these procedures. Curriculum elements are defined to provide a broad basis for delineation of useful evaluative foci. The social context of evaluation is recognized by provision of the "paradigmatic criteria" dimensions. Juxtaposition of these criteria (evaluator's and actor's) suggests the means for delineation of issues.

Obtaining of information is indicated in the model by analysis, interpretation and judgement (against the evaluator's paradigmatic criteria) of observations.

Providing of information to decision-makers is accomplished in the model by a process of "qualitative exposition" which was defined as a richly descriptive account of the evaluator's findings.

The model for this study also possesses sufficient flexibility to accommodate various evaluative functions regarded as desirable or necessary by recent evaluation theorists, i.e. "formative" and summative" assessments (Scriven, 1968), evaluation for course improvement (Cronbach, 1963; Scriven, 1968), assessment of curriculum materials (Eash,



1970; Morrissett and Stevens, 1967; Sjogren, 1970: cited Worthen and Sanders, 1973).

A survey of existing evaluation models provides for the reader a large variety of direct prescriptions for the conduct of evaluation, and indirect conceptions of the nature of curriculum and curriculum development. Evaluative models that are designed to play a role in course improvement ( and most are) demonstrate an intimate relationship between evaluative processes and curriculum development. Different models define and arrange components of curriculum in diverse ways with varying degrees of specificity. The model proposed in this study accommodates under its six previously defined elements, all components of curriculum identified by the other models reviewed in Chapter II.

Integration of the curriculum evaluation model for the visual arts with existing theories of evaluation can therefore be claimed in that:

i) it demonstrates a similar basic definition of process of evaluation possessed by other models.

ii) it is designed to perform similar roles to other models (i.e. formative and summative).

iii) it encompasses within its structure definitions of the components of curriculum delineated (or implied) by other models.

The second question asked by the norm of coherence is: In what respects does the model differ from existing models?

The differences between the model of this study and





existing models of evaluation have been stated clearly at various points in preceding chapters. Within the statement of Purposes of the Study in Chapter I the necessary characteristics of an evaluation model appropriate to the visual arts were listed. These three characteristics represent a succinct description of the essential differences between this model and others. Following is a brief description of the meaning of each.

The first characteristic possessed by the curriculum evaluation model for the visual arts is an open frame of reference. Other models assume a linear type of design. This means that neither evaluative nor curriculum development sequences are pre-determined in the "art" model.

Most models of evaluation assume the "logical" course of behaviour in curriculum development to be, in very general terms: i) assessment of needs and perhaps system capabilities (reasons); ii) determination of rationale and broad objectives; iii) derivation of specific objectives from the general statements; iv) arrangement of objectives hierarchically; v) selection of the means to achieve the objectives; vi) implementation of the programme. As most evaluation models are concerned with contributing towards course improvement their evaluation processes are inextricably linked with the "logical" sequence of curriculum development.

Models which are not connected with the development process (i.e. are concerned only with outcome assessments) reflect to a greater or lesser degree the Tylerian approach. This type of evaluation characteristically observes the



following sequence: i) discovery of the programme's objectives; ii) construction of operational definitions of desired outcome behaviours; iii) selection or development of appropriate instruments to measure those behaviours; iv) administration of the instrumentation; v) comparison of the observed with expected outcomes.

It is recognised that the above analyses are very general and that many variations and refinements have been introduced within different models. Also, the existence of valutive procedures for judging the worth of objectives, content, and instrumentation in some particular models is acknowledged. However, despite variations, in general terms the previously described "logical" process of curriculum development and evaluation exist as an underlying assumption of current evaluation models.

The same assumptions of "logical" procedures for construction of curriculum and conduct of evaluation do not underlie the model proposed for the visual arts. This is not to say that this model is built upon assumptions of illogical or irrational conceptions of curriculum or evaluation. Rather, the open framework accepts that the means and ends of art education can, and usually do, exist in an interactional relationship and are not clearly separable as in the classical Tylerian mode. In other words, art programmes tend to proceed in a purposive rather than a purposeful fashion.

The major implication of the means - ends interaction for curriculum development and evaluation in the arts is that a clear picture of the end product (i.e.



individuals who are educated in the arts) is not available to educators or evaluators at the beginning of the programme.

The processes of curriculum development, instruction and evaluation are all affected by this circumstance. Art programmes, for example, may be developed by defining broad areas for exploration, then developing activities judged to provide potential for worthwhile experiences within those content areas.

Instruction may proceed from topic to topic, or activity to activity within the broad context defined by the curriculum but particularly defined by student needs, interests, unexpected outcomes, the teacher's area of expertise, and so on. In short, goals can be suspended in favour of attending to the ambivalent qualities of the immediate situation (Aoki, 1973). The evaluation question becomes, in this kind of situation, not so much whether the means effectively achieve the desired pre-specified ends, but that the experiences gained along the way were intrinsically worthwhile. Stake (1975) made this observation:

Many of my fellow evaluators are committed to the idea that good education rests in specific outcomes: performance, mastery, ability, attitude. But I believe it is not always best to think of the instrumental value of education as a basis for evaluating it. The value may be diffuse, long delayed, or forever beyond the scrutiny of evaluators. In art education it is sometimes the purpose of the program staff, or parent, to provide artistic experiences and training for the intrinsic value above. "We do these things because they are good things to do," says a ballet teacher. The evaluator or his observers may form opinions as to whether or not these things were done well... He should not presume that only measurable outcomes testify to the worth of the program. (p. 16)





This is not to suggest that nothing measurable emerges from an art programme. On the contrary, many skills (particular productive skills) and concepts are testable. Information gained through their measurement may also provide useful information with respect to inferences about the success of certain aspects of a programme. However, measurement designed to assess the degree of attainment of pre-specified behavioural objectives is not considered by many art educators to be the central purpose of evaluation of art education (Bradley, 1969; Efland, 1974; Eisner, 1973; Stewart, 1972; Kaelin, 1969).

The framework of the evaluation model for the visual arts is open (in comparison with other models) in that sequences of curriculum development activities are not prescribed. In addition "worthwhileness" of educational outcomes may be assessed in this model through application of the evaluator's "paradigmatic criteria" rather than comparing specific objectives with observed (measured) outcomes.

Use of this framework, it should be emphasized, does not preclude a possible role for the "logical" model. It simply provides the evaluator with the opportunity to exercise appropriate choices where necessary.

The second characteristic possessed by the "art" model is closely related to the first, in that the "open framework" discussed previously implies use of judgemental and interpretive methodologies rather than, or in addition to measurement technology. The significance of this, in terms of contrast with other models, lies in the varying





conceptions of validity employed by alternate epistemologies. The validity issue was explored at length in Chapter IV, so will not be pursued here.

The third characteristic of the "art" model is related to the second. Use of judgemental and interpretive techniques allows preservation of individual or small groups of behaviours in the analysis of data. Particularly significant events may be chosen by the evaluator for highlighting, rather than to be lost in generalizable information. This kind of procedure is necessary to assist the evaluator in his preparation of the evaluative product - the "qualitative exposition".

The third question asked by the norm of coherence is: Is the logic of the model internally consistent? In the context of the proper scientific use of theory (or model), this is a necessary question for validation of scientific explanations. However, in the sense that "model" is used in this study (i.e. as a set of propositions for action, rather than as an explanation of phenomena) the question is rather more difficult to answer. The root of the difficulty stems from the fact that the meaning of "evaluation" is socially constructed, rather than empirically derived through deductive methods. The question of logical consistency is therefore not demonstrable in any satisfactory way. The only statement which may serve tangentially to satisfy the question is that there are no obvious contradictions existing amongst the set of propositions which constitutes the model.



### Pragmatic Norms

Three questions are posed by this norm: "How well does the model function?"; "Does it serve evaluative purposes?"; and "What is the nature and usefulness of the information generated by the unique characteristics of the model?"

These questions are dealt with by presenting a brief report of an evaluation conducted for the Edmonton Public School Board. This evaluation occurred in connection with the development of a set of Art Appreciation Materials. The report is interspersed with discussions designed to illustrate methodological procedures used, as well as to demonstrate their utility.

### Evaluation of the EPSB Art Appreciation Project

#### Background to the Evaluation

A brief description of the project was provided in Chapter V. As was mentioned in that description, the primary purpose of the project was to develop a set of curriculum guide materials in Art Appreciation, to serve as a supplement to the Alberta Elementary Art Curriculum Guide. The project was initiated in March 1974. The writer of this study was associated with the project as external evaluator from October 1974 to June 1976.



During that time the project passed through two major stages. The first was development and refinement of a unit of materials called Introducing Ways of Seeing. This unit represented the first of four related units planned for completion in 1977 or 1978. The second major stage of the project was pilot testing of the first unit of materials in 30 schools.

The initial writing of Introducing Ways of Seeing was done during the summer of 1974. Refinement of this material occurred during the winter of 1974-75. During the summer of 1975 re-writing and expansion of the first unit was undertaken.

During the winter of 1975-76 Ways of Seeing was pilot tested in 30 classrooms. Teachers were selected for this pilot on a volunteer, first come - first served basis.

The evaluative function during both stages was essentially formative. During the first stage the primary evaluative contribution was a content assessment of the preliminary material. To effect this evaluation the writer attended all meetings of the development committee during the winter of 1974-75. In addition, individual interviews of each of the committee members were conducted. The purpose of this action was to discover the values (and paradigmatic criteria) held by the members regarding the project. This information assisted in developing appropriate questions for content analysis. A questionnaire



was then constructed.

A sample of potential users of the materials was selected (41 elementary teachers). Copies of the initial materials, and the questionnaire were then distributed to these teachers. These respondents supplied data relative to the utility of the material, clarity of its structure, concepts, and format, and the need for such materials in their schools.

Also during this time the prototype of the evaluation model was being developed. The questionnaire format used for the content analysis reflects, to a certain extent, the structure of that prototype.

The results of the content analysis were submitted to the EPSB committee on June 30, 1975. Recommendations for refinement of the material were observed by the writers in their revisions.

The major function of the evaluator during the second stage of the project (wider pilot testing of the revised materials in 1975-76) was to assess the potential value of the programme in the classroom context. At this time the revised version of the evaluation model had been developed. The evaluation design and methodology used for this assessment was based upon the structure of the revised model.

The procedures used to conduct the evaluation of Stage I of the project are not of direct importance for this study so will not receive further elaboration.





The Stage II evaluation bears direct relationship to the study and is discussed in greater detail below.

### Purposes of the Evaluation

The major purpose of the Stage II evaluation was to determine the value of Ways of Seeing within the classroom context. Essentially, as the assessment was conducted in a pilot-testing situation, the evaluation question was one of assessing the potential value of the materials.

A number of sub-questions were derived from this major one. These were generated through the evaluator's interaction with the committee, from analysis of the results of the earlier Stage I evaluation, and from observation of the materials in use. It should be noted at this point that the questions were not posed prior to the evaluation but arose from discussions and observations conducted during the course of the study.

The format adopted in this study arises from reasons of convenient presentation only.

Each of the evaluation questions stated below is accompanied by a brief justification of its value.

1. How valuable were the outcomes of observed lessons in terms of quality of the experiences gained by the students?

This question was asked in lieu of its counterpart (in more traditional evaluation designs) which could be:



Did the observed lessons achieve their stated objectives? Such a question was inept in this particular situation due to the nature of the objectives that were stated, and the overall philosophy of the programme. The broad objectives for the entire Ways of Seeing unit are presented in Appendix "D". Some selected examples of specific lesson foci (objectives), together with their related lesson activities, are presented in Appendix "E".

The reader will appreciate that the nature of these objectives and foci are such that the types of experiences and investigations to be encountered by students are outlined, whereas specific learning outcomes are not. This kind of format was deliberately employed by the developers to encourage attention to the open ended nature of artistic experiences. Specification of particular outcomes was avoided so that lesson units would not proceed in a "closed" linear fashion.

Value of observed lessons could only be determined in this instance by reflective analysis, rather than by establishing beforehand the necessary conditions for success.

The question itself is the most significant in terms of the overall evaluation. The curriculum materials may be judged to possess many good qualities but the ultimate test of their worth must inevitably be assessed on the basis of classroom success.



2. How well are "non-art" elementary teachers able to direct art appreciation experiences, given the open format of the guide materials?

A "non-art" teacher is defined as a teacher with fewer than the required number of courses for a "minor" specialization in art. To participate in "art appreciation" one must exercise discriminatory perceptions. It is logical to assume that teachers charged with the responsibility of directing students' aesthetic experiences should themselves be capable of making fine discriminations. This is particularly significant in light of the "non-specificity" of outcomes characteristic of the curriculum materials. Other questions related to this issue are: "How do non-art teachers judge the success of art appreciation lessons?" "How well do non-art teachers direct discussions arising from the suggested art appreciation activities?"

This question is important in light of the fact that these guide materials were developed specifically for non-art teachers. One of the shared paradigmatic criteria of the curriculum developers, frequently expressed at committee meetings, was that the guide materials should be designed so that non-art teachers would experience no difficulty in their use. In fact, it was also hoped that use of these materials would assist in development of the teacher's perceptions as well as the students'.

3. Do the pilot teachers use the curriculum materials



as intended by the developers?

Misunderstanding of the nature and purpose of curriculum materials can lead to their inappropriate use. On occasion positive benefits may be derived from unexpected, imaginative interpretation of guide materials by teachers. This question was considered, therefore, to be important in terms of its analytical value. The following statements indicate the intentions held by the developers for use of the materials. These statements were derived from the Rationale and Explanation of the

- Format:
- i) The primary purpose of the materials is to assist in refinement of visual awareness.
  - ii) Although the exercises are activity oriented, products are not considered as the most important outcomes of experiences.
  - iii) Appreciation exercises should not be restricted to the study of art works.
  - iv) The content should be easily correlated with other subject areas.
  - v) The programme is not necessarily the whole elementary art course - it may be a major or minor supplement to it.
  - vi) Activities are not necessarily designed to be approached consecutively; rather, teachers should select from the range of activities those that are appropriate to their needs.
  - vii) Time required for specific exercises is





variable.

viii) Language should be taught where required rather than as the central purpose of specific exercises.

4. How useful are the curriculum materials, in the eyes of the pilot teachers?

This question was posed as a consequence of the findings of the Stage I evaluation which revealed some vagueness of terminology and clarity in the materials. It was anticipated that such vagueness, if not corrected, would reduce the usefulness of the materials. Rewriting occurring between Stages I and II was intended to correct this problem.

The content analysis of the Stage I evaluation was conducted by potential users of the materials, not by teachers who had tried them. Consequently, questions of content, clarity, structure and terminology were deemed worthy of exploration in the pilot situation.

5. How does availability of visual aids affect the programme?

Any art appreciation programme obviously requires some "art" to "appreciate". Numerous visual and audio-visual resources are listed in the preliminary draft of Ways of Seeing. The Stage I evaluation indicated that these aids were difficult to get, particularly from the EPSB Instructional Media Centre. A logical follow-up question for the pilot study, therefore, should be related



to the availability and function of aids.

6. How useful were the introductory meetings of the pilot group?

Problems of dissemination and implementation often work against effectiveness of new programmes. The purpose of this question was to focus upon the implementation procedure employed.

Each of the above questions represents a focus of the evaluation. The elements of the evaluation model proved useful as a guide for directing evaluative attention to general areas of curriculum to search out possible questions. Table 15 indicates the relationship of the evaluation question to the elements of the model.

#### Methodologies Employed for the Evaluation

In keeping with the open framework of the model, evaluation issues and questions were not pre-determined. In the initial stages, while working with the committee, observations and informal discussions were the primary modes of data collection. The kinds of procedures used were analogous to the methods advocated by Stake (1975) and Schatzman and Strauss (1973). During the first classroom visits general procedures and routines of teachers and students were noted. The evaluator acted as an observer in a similar manner to Smith in the Smith and Geoffrey (1968) ethnographic study.

The guiding research interest of the initial



TABLE 15

Relationship of Evaluation Questions andModel Elements

Question	Model Elements
How valuable were the outcomes of observed lessons in terms of quality of the experiences gained by the students?	Outcomes
How well are "non-art" elementary teachers able to direct art appreciation experiences, given the open format of the guide materials?	Agents
Do the pilot teachers use the curriculum materials as intended by the developers?	Rationale Implementation
How useful are the curriculum materials, in the eyes of the pilot teachers?	Content
How does availability of visual aids affect the programme?	Resources
How useful were the introductory meetings of the pilot group?	Implementation



observations was also the overarching purpose of the pilot (and the evaluation); to discover the potential value of the curriculum guide materials. As issues began to emerge and questions became clarified, research methods were selected that were judged most suitable for the purposes at hand.

"Within Classroom" Methodology. As was mentioned previously, Eisner's (1975) concept of educational connoisseurship was tested as an "interpretive" evaluation methodology in the classroom situation. Connoisseurship may be regarded as an appreciative art requiring "refined" perceptions. Exercise of this refined perception in an art context is employed for the purpose of gaining understanding of meaning. Achievement of understanding subsequently supports judgemental (or critical) activity. The important difference between criticism of two dimensional art works and curriculum activity was discussed previously.

Translated into the classroom context "connoisseurship" appeared to offer the researcher a choice between two basic observational methodologies. These were either procedures of "non-interactive" observation that employ the use of mechanical or manual recording devices, or "interactive" methods which use informal conversational and interview techniques.

The "non-interactive" procedures in the classroom situation provide data for inference. For example: "Ron swung back on his chair and stared around the room instead





of getting on with the exercise"; is an observation that could have produced the inference that Ron was obviously bored with the exercise. In fact "interactive" procedures (discussion with Ron) could reveal that Ron had to share a pair of scissors with Dwayne and was simply waiting his turn. Moreover, he had a multitude of ideas for the exercise.

As the study progressed, more useful interpretive information was found to be gained through interaction with the actors than through external observations. This information was regarded as useful because it allowed the researcher to understand the meanings attributed by the actors to their actions and the events resulting from programme activities.

Both methodologies were employed for the purpose of developing as full a picture as possible of the classrooms involved. In light of the discussion of validity in Chapter IV it is important to distinguish clearly between the kinds of information sought through "interactive" and "non-interactive" observational methodologies.

The construction of meaning, through informal discussions, and questioning actors about the reasons for their actions, produces explanations of the actor's view of the classroom world. "Interactive" information gained by a researcher is useful in that he is able to gain an understanding of the "multiple realities" that exist



within a single classroom context.

Meaning ascribed to "non-interactive" data, on the other hand, is a statement of the researcher's view of the classroom "reality". The researcher is able to ascribe meaning only in terms of his own previous experience (immediate or vicarious) of "similar" situations if he does not have the benefit of interaction with the actors he observes.

Two different notions of validity also apply in relation to each type of data. "Interactive" information is gained through methods akin to phenomenological research. The reader will recall that the primary validation criteria for phenomenological explanation is the postulate of adequacy (Schutz, 1973). Therefore, the researcher must demonstrate that his explanations of actions are recognized by the actors.

Explanations derived from "non-interactive" data in the interpretive model for art curricula represent judgemental statements (inferences) made by the evaluator. Equivalent judgements made by a critic in the artistic mode of inquiry must possess referential adequacy to be valid. In the classroom context the researcher should ensure that his interpretations of "non-interactive" data are also referentially adequate.

It was recognised that the two procedures, "interactive", and "non-interactive" would affect each other when both were used in classroom observation. As



the evaluator gained understanding of the actors' meanings and values it is likely that somewhat different judgements were made from "external" data. Following from these judgements, different avenues of questioning were explored with the actors.

The ultimate aims of these activities in the EPSB evaluation were to derive sufficient data for the evaluator to make judgements about the value of observed classroom events, and to identify issues that may have required clarification outside the classroom. In types of evaluations (other than the EPSB project) it is recognized that data different from "interpretive" observational information may be useful or necessary.

Other methodological issues are raised by the above discussion. The question of observer effect is a perennial source of dispute in any discussion of research methodology. Matson's (1964) point that observer interference is unavoidable in even the most exact of the natural sciences is useful to note in any reference to the issue. The implication for social science research is that no matter what steps the investigator takes to control the effect of his presence he cannot avoid affecting the "natural" course of events. The extent of his influence is impossible to assess in any event as the completely "natural" situation lies outside the range of research experience. The only possible approach in the classroom is to admit some effect while attempting



to tread the fine line which produces the most useful information with minimal intrusion.

"Outside Classroom" Methodology. Most of the valuable data collected outside the classroom were "interactive" in nature. Informal discussion with committee members, and teacher interviews constituted the primary research style. Considerable initial observation preceded much of the interactive work. The Schutz postulate of adequacy was again employed as the main criterion for validity. For example, the researcher repeated key statements recorded in note form during teacher interviews to ensure that the "reduced form" held the same meaning for the interviewee as his previous "long" explanation.

Validation. The interpretive framework carries with it a different philosophical framework regarding validity of its findings than does the traditional evaluation model derived from the empirical-analytical tradition. As in art criticism an evaluation using the proposed curriculum evaluation model for the visual arts represents a judgement supported by evidence from the object of investigation. An evaluation supported by empirical-analytical findings claims to represent the "true" state of affairs or the "true" effects of the programme.

The interpretive framework does not claim "truth" of judgement in any absolute sense. Rather there





are only judgements which are more or less strongly supported by evidence. The judgement represents a statement of the evaluator's interpretation of the "reality" of the curriculum. His "reality" however is formed by attending to the "multiple realities" that exist within the social context of any curriculum situation.

The empirical-analytical claims to "truth" or "accuracy of findings" are questionable in light of the discussions of validity in Chapter IV. The same arguments used by the empiricists with respect to the likelihood of differences between opinions of judges (particularly in an art context) can also be turned around and levelled at the empirical-analytical tradition. Katahn and Koplin's (1967) observations regarding the paradigm clash which exists within the field of psychology well illustrate this point. These authors attributed the apparent clash between behaviouristically oriented psychologists and those who prefer cognitive interpretations to a paradigm difference.

In these controversies each of the antagonists describes his opponent's position in the most extreme form in order to formulate telling counter arguments. At the same time each insists that the other's criticisms are based upon inadequate understanding of one's own preferred theoretical framework as well as a lack of knowledge concerning the relevant supporting data (p. 147).

With respect to the initial choice of which evaluation questions to attend to, there is very little difference between the empirical-analytical scientist



and the "educational critic" (as was pointed out by Mann, 1969). It is not an unreasonable suggestion in light of the preceding discussion that two "empirical" evaluators may derive operational definitions and choose, or develop, instrumentation to assess the EPSB project, yet arrive at varying conclusions. This variance would be attributable to differences in the researchers' basic philosophic orientations.

In the art appreciation project the writer was the only classroom observer and interviewer. Limited use was made of another observer to view video-tapes. That additional opinions would have added substance to the evaluation is undeniable. With a larger number of expert opinions to choose from, the evaluator would be able to support his judgements upon a wider data base. It is noted elsewhere in this study, however, that the ultimate choice about which observer's judgements are the most strongly supported (in the event of a dispute) should rest with the evaluator. This issue is discussed more fully under the next section - Pragmatic Value of the Model.

The appropriate validity checks, discussed in the previous section, were made by the researcher throughout the study. Therefore, despite the disadvantage of limited manpower, it was felt that the findings of the study possessed sufficient inherent verification to serve evaluative purposes. In other words, the researcher, through constantly referencing his findings against the



"object" of investigation and the actors' perceptions, satisfied himself that his observations were adequate.

### Design of the Evaluation

Procedure. The design of the EPSB evaluation study was implicit within the preceding discussion of methodology. In more clearly defined terms, the procedure of the evaluation was as follows:

1. Observation of monthly curriculum development committee meetings from October, 1974 to June, 1976.
2. Informal discussions with committee members at random intervals during that time.
3. Preparation and distribution of questionnaires to 41 potential users of the curriculum materials.  
(Findings from this questionnaire comprised the Stage I evaluation report submitted to the committee June 30, 1975).
4. Observation in the selected pilot classrooms from February to April, 1976 (inclusive). Details of classrooms are provided below.
5. Discussion and interview with students and teachers in the pilot classrooms.
6. Definition of evaluation questions emerging from procedures observed in 1-5 above.
7. Discussion with project director and assistant to verify shared understanding of the significance of evaluation questions.
8. Informal discussion and interview with 21 pilot



teachers.

9. Preparation and submission of findings.

As was noted in the discussion of methodology, two observational styles ("interactive" and "non-interactive") were used. The "non-interactive" mode was used predominantly in the early stages while the "interactive" was increasingly employed as issues became defined during the later stages.

The two general data sources were defined as "within classroom" and "external to the classroom". Table 16 shows the relationship of the type of data collected and their sources.

Classrooms. Two classrooms were selected for extended observation. Three teachers volunteered to assist with the evaluation. One taught grade four and was a writer of the materials while the other two, team taught grade six in an open-area classroom.

As the overall purpose of the evaluation was to determine potential value of the curriculum materials, volunteer selection was considered appropriate. The grade four teacher possessed an extensive art background while the two grade six teachers had non-art backgrounds. The contrast of grade level, situation, and teacher background was viewed positively in that it would provide opportunity for observation of different interpretations of the materials.

Observations. Regular visits were made to schools during the months of February, March, and April, 1976.





TABLE 16

Relationship of Data Type and Sources

Source		Data Types
General	Particular	
Outside Classroom	Monthly meetings of EPSB Curriculum Development Committee	Field notes
	Teacher Interviews at Pilot Schools	Field notes Interview Schedule
Within Classroom	Grade 4 Conventional Self contained classroom	Field notes (from formal and informal discussions) Audio Tapes Video Tapes Photographs
	Grade 6 (Two) Open area-team taught classes.	Copies (Xerox) of student products



These visits generally occurred on the "art" day although in the initial stages visits on non-art days were made. The purposes of these visits were to allow students and teachers to become familiar with the presence of the evaluator and to permit the researcher to familiarize himself with the normal classroom procedures. Later visits either began before the art lesson itself or extended beyond it where possible. Pertinent information with respect to classes, visitations, and lessons observed appear in Tables 17 and 18. Table 17 refers to school A, and Table 18 refers to school B.

Interviews. Of the 30 original pilot teachers three refused interview on the grounds that they had not used the guide materials and consequently could not assist the researcher. Reasons for not using the material varied. One had not received a copy, one had changed her mind about participating in the pilot, and the third suggested that too much work was required to implement the programme.

Interview could not be arranged with five other teachers for various reasons. Of the remaining 22 pilot teachers, 21 were interviewed. One was absent on the day of the scheduled interview.

### Derivation of the Findings

Following is a discussion which provides a description of the circumstances that led to formulation of some of the evaluation questions. As well, examples



TABLE 17  
Class A - Visitations and Lessons Observed

Visit No.	Date	Length of Visit	Lessons
1	Feb. 4	2½ hours	Stitchery (Value, Colour, Texture, Unity)
2	Feb. 6	2½ hours	Music, Reading, Spelling
3	Feb. 18	2½ hours	Exploring Colour
4	Feb. 25	2½ hours	Exploring Colour
5	March 3	2½ hours	Exploring Colour
6	March 12	2½ hours	Exploring Colour *
7	March 16	2½ hours	Exploring Colour **
8	March 24	2½ hours	Exploring Colour Printers Ink *** and Brayers
9	April 9	2½ hours	Exploring Colour Printers Ink ** and Brayers

Student No. = 27

\* Minor Variation upon suggested activity

\*\* Extension of idea from previous lesson.

\*\*\* Additional activity congruent with the focus (i.e. exploring colour)



TABLE 18

Class B - Visitations and Lessons Observed

Visit No.	Date	Length of Visit	Lessons
1	Feb. 13	5 hours	Reading, Math, Art, Social Studies
2	Feb. 20	2 hours	Exploring Colour *
3	March 4	1½ hours	Exploring Colour (Continued)
4	March 19	1½ hours	Object Drawing **
5	March 26	1½ hours	Object Drawing (Continued) **
6	April 23	1½ hours	Exploring Texture
7	April 30	1½ hours	Exploring Texture (Continued)

Student No. = 44

\* Minor Variation upon suggested activity

\*\* Major variation upon suggested activity





of the findings and samples of the evidence which supports them are used to illustrate the "open" nature of the research procedure used.

The first evaluation question, as was mentioned before, was the most important and guided initial observations made in the two situations. This question was: How valuable were the outcomes of observed lessons in terms of quality of the experiences gained by the students?

In order to make judgements with respect to this question the researcher acknowledged that curriculum materials alone would not be responsible for good or poor lesson outcomes. The teacher was recognized as a critical determining variable. Therefore in focussing upon experiences gained by students as a result of engaging in the activities suggested by the materials, some attention was also paid to the ways in which teachers interpreted and implemented the programme. To provide the fullest indications of teachers' "styles" of using the materials, all lessons, directly or indirectly related to the content were observed. This course of action was considered to be most productive as the potential value of the materials could only be assessed in terms of possible teacher interpretations and uses.

The first significant finding was that all observed lessons provided some valuable experiences for the students. Highlights of the positive experiences are



reflected in the following examples from the data:

i) School A - Visit 4. The lesson was based directly upon Focus 6, Activity 1, presented in Appendix "E". Students dressed in blue, constructed eye glasses from blue gels, painted with values of blue paint, and listened to "blue" music. Students were stimulated by the novelty of the classroom atmosphere. (2" x 2" colour slides were collected for the EPSB report, to recapitulate the classroom mood).

Student discussions following the lessons revealed highly personal, diverse, and sometimes penetrating observations regarding the mood produced by the colour blue. Here are some samples from a transcript of the lesson.

Teacher: How do you feel now that you have been blue for about one and a half hours?

Kim: I feel creepy.

Susan: I get a bored feeling because everyone is dressed the same. It is monotonous.

The following responses were produced by the teacher's question, "What does blue remind you of?":

Denise: Blue reminds me of myself.

Teacher: It reminds you of yourself? How does it remind you of yourself?

Denise: Whenever I go to bed I always think of blue.

Teacher: Why?

Denise: It helps me go to sleep.



Teacher: It makes you feel rested?

Denise: Yes.

Kevin: I think that Heaven would be blue.

Tessy: Um - like when you're on the sea-shore, like, and it rained in the afternoon and the wind is blowing, the cool wind, and you're walking with your bare feet in the sand.

Teacher: Is the mood blue or are you talking about the sea blue?

Tessy: Everything feels blue.

The discussion ranged from students' explanations of the significance of blue in their designs and paintings, through the naming of things that are blue, blue smells, speculations about the effect of changing the colours of things (e.g. blue mashed potatoes) to the teacher's reading of a poem about blue.

This demonstrated a marked contrast to the first lesson observed, in which the students were required to select colours for a stitchery project. Reasons offered by students with respect to their choices of colour were restricted to like/dislike, and realistic considerations (e.g. "I chose grey for the rocks because rocks are grey" - Kim). The "blue" lesson was therefore judged to represent a positive expansion of students' experience of colour.

ii) School A - Visit 6. This was a colour mixing lesson. In a prior lesson students viewed the film Discovering Colour and learned the names and meaning of



the terms "primary colour", "secondary colour", and "value". The observed exercise required students to write themselves a contract for a colour mixing exploration, then to carry out the contract. (Samples of the contracts were submitted to the EPSB Committee with the evaluation report).

The students became engrossed in their colour mixing activities; few of the contracts were the same. During the course of the lesson the researcher was able to interview 15 of the class members. All but one were purposefully engaged. Of those 14 students, most had made a variety of discoveries. The following excerpts from the taped interview illustrates the kinds of discoveries students achieved:

Researcher: Tell me what your're doing, Scott.

Scott: I'm adding the primary and secondary colours and mixing two colours or three colours together and trying to figure out what kind of colours I'm making.

Researcher: Tell me about some of the colours you already have.

Scott: That one is green, and that one is orange and this one was blue and purple, and this one was a red and blue. This one was red and orange, a bit of brown and yellow.

Researcher: Did they all turn out the way you expected? Or didn't you know what to expect?

Scott: I didn't know what they were going





to turn out to.

Researcher: Were any of them a surprise to you?

Scott: That one.

Researcher: Which one was that?

Scott: The one with the orange and red  
and yellow and a bit of brown.

Researcher: What did you think you might have  
gotten when you did that?

Scott: Black.

Researcher: Lisa, tell me what you're doing there.

Lisa: I'm making different shades and tints  
out of blue.

Researcher: I see, is it working out like you  
expected it to work out?

Lisa: Not really.

Researcher: What's happening?

Lisa: Well I thought that would work out,  
like, ... you'd see more blue than black ... like,  
you see a whole bunch of black instead of blue.

Researcher: Did you find out anything about black?

Lisa: It makes colours darker.

Researcher: Kim, would you like to tell me about  
your work?

Kim: I'm trying to make it real light and  
keep on getting lighter and darker and darker and  
darker, so that it goes from lighter to dark.

Researcher: What do you call that?



Kim: Well I just call it adding white and black.

Researcher: Were any of those a surprise to you - those colours?

Kim: Yes - this one - its the colour of my room!

The follow-up discussion to this lesson proved valuable in that the students were able to share different discoveries. Also, the teacher pursued the "unexpected" outcomes to fruitful conclusions. For example, several students did not achieve the colours they had expected. Questioning and explanation by the teacher revealed that some students were using disproportionate amounts of paint, that some colours were stronger than others, and some were allowing their first colour to dry before adding the next, and consequently unexpected results were produced.

iii) Class B - Visit 3. This lesson was very similar in intent to the one described previously for Class A. Students viewed two films about colour mixing, one related to paint, the other to light. The problem posed for students was to plan a means to demonstrate the mix of paint or light as demonstrated in either of the films.

The discussion of the outcomes of this exercise provided particularly valuable results in the judgement of the evaluator. A very wide variety of means was employed by the students to demonstrate ways of mixing



paint or light. (A set of 2" x 2" slides was included in the EPSB report to demonstrate this variety).

Discussion by the students reflected their strong interest in the task. Animated discussion occurred on several occasions in attempts to explain unexpected results.

Following is an example from a transcript of the lesson:

Doug: First I started by taking the primary colours to make the secondaries. This one didn't quite turn out (pointing) but it turned out a red with streaks of yellow - not really an orange. There's a bit of purple (pointing to a second colour). And this one's a black, green-blue. It's not even really a turquoise colour either.

Then I used orange with the three primaries and then I used purple with the three, then green with the three. Then I would combine the secondaries, the three secondaries, just like I had done with the primaries. And then all I got were these oranges. This one turned out like I expected, a light orange colour.

This turned out somewhat like I expected even though I put yellow on top, it turned out more yellow than I wanted. This one turned out really odd. I have no explanation for why it turned out black.

Teacher: I think you had some change in your plan while you were doing that with the



kind of paper you were using.

Doug: Yes I started out with normal everyday paper, then I thought it would combine better if I had this slippery type of paper, this chart type paper.

Teacher: What was happening while you were using more coarse paper?

Doug: The colours weren't combining.

Teacher: O.K. are there any suggestions then as to how Doug might get a truer mixing of colours without them soaking into the paper?

Sherrie: Mix the colours on something the paint wouldn't soak into.

Teacher: What kind of thing might that be?

Sherrie: A piece of plastic.

Cam: Metal.

Teacher: What would happen on wood.

Danielle: It would soak in.

Teacher: Probably.

Danielle: On plastic it wouldn't go on right, like it would go on in globs, it wouldn't go on all over.

Sandra: I tried mixing oil paints. I asked for oil paints. I don't know if they're true oil paints or not but they wouldn't mix together.

Doug: I wonder what kind of paint they used on that film because it sure wasn't this kind.





Dawn: I used opaque water colour and it worked really well.

Although similar kinds of successful results were achieved by the colour mixing lessons within both classes the teachers' reactions regarding their success were different. The teacher in Classroom A felt her lesson was valuable because the students were able to formulate and pursue their own problems, many problems were pursued at the same time, and students were able to learn from each other. The teacher also felt that the lesson served to demonstrate which students were having trouble with the meanings of some of the technical terms.

Both teachers in Classroom B were disappointed with the outcomes of their colour mixing lesson. Both felt the results were too uncontrolled due to too much leniency allowed with the media, and the outcomes were consequently too diverse. In fact, those teachers ranked the lesson amongst the least successful they had tried.

This difference in viewpoint between the teachers led the evaluator to re-examine the two approaches that were reflected in the teacher's implementation of the lessons. In Classroom A the task was one of exploration. Students were requested to state in their contracts the specific nature of their exploratory task, not the expected outcome. The diverse results subsequently obtained were welcomed by the teacher and shared by the students.



In Classroom B the students were required to pre-determine their desired result (i.e. to demonstrate a particular colour mixing fact), then select appropriate means to proceed towards that goal. This was particularly evident during the follow-up discussion in Classroom B. Students were requested to report their "plan" and explain the reasons for any deviation from it.

The essential difference between the two approaches, despite the fact that similar results ensued, was that Classroom teacher A adopted an "open ended" approach whereas Classroom B teachers used a linear procedure.

Other manifestations of this difference between the approaches of teachers in both situations were noted. In Classroom A, a student commented that the colour blue reminded her of heat: "Of the sun's rays coming all the way down to the earth". This observation was accepted by the teacher who returned to it later (in a language arts lesson) following the reading of a poem about blue. In the poem a reference was made to the "hot, wild, screaming blue" of a welding torch. The teacher explained that although the colour blue is normally associated with coolness, occasionally it reminded people of heat.

In Classroom B, on visit 5, the students were asked to name "the" warm and "the" cool colours as though this information constituted "fact". The possibility of alternate responses was not considered.

During the entire period of observation in



Classroom B not a single instance of a student's personal response to (feelings about) a work was recorded. All discussions were technical or "fact" related. In Classroom A, on the other hand, strong emphasis was placed upon students' feelings about visual phenomena.

Another significant incident illustrating the linear, objective, and "factual" approach of Classroom B occurred during the object drawing lesson (Visit 4). Students were requested to bring a simple object from home to draw. The objective of the lesson was to make as realistic a drawing as possible using knowledge about colour and value to achieve the illusion of three dimensions.

The following exchange was taped as the researcher discussed a student's drawing with him. During the discussion one of the teachers joined in to talk about the student's work. (The student was drawing a model aircraft he had built himself):

Researcher: What's the aim of the lesson, Richard?

Richard: We're supposed to be looking at colour and texture of different things.

Researcher: Okay.

Richard: And shading. Like really, when you look at things you get different colour and there is shading on it. Like say if you got an angle like this (demonstrated with model aircraft) you have shadows here (pointing).

Researcher: What are you drawing, is it a



Spitfire?

Richard: No it's a Mustang.

Researcher: Why did you choose it?

Richard: I like drawing aeroplanes. I have a couple at home and this was the only one I could bring.

(At this point the teacher who had overheard the discussion broke in).

Teacher: You were asked to bring a simple object, that's not.

Richard: Well no, not really.

Teacher: Tell me, which way are you looking at that?

Richard: Well I'm looking at it like this, (student held up model so that it was viewed from the top).

Teacher: Do you want to look at it like that? Or do you want to look at it like this?

(Teacher indicated side view of model).

Richard: Like this. (Student indicated previous view of model).

Teacher: Why?

Richard: Well like this it looks as if it's coming round into a dive. (Student indicated the action of a fighter banking prior to diving).

Teacher: Which is the most pleasant view to the eye? Which angle does it look best at?





What is the most pleasant looking angle to the eye?

Richard: Well, sometimes it's like this.

(Student indicated top view). There's a lot of angles.

Teacher: Something like this? On a slight angle but not completely flat - all right?

(Teacher indicated three-quarter view of the model).

Richard: All right but I want it like that.

(Student indicated the same view as before).

Teacher: Now do you think you could sketch something like that? (Three-quarter view indicated).

Richard: Yes.

Teacher: Right, well try something like that. It is more interesting to look at. It is more...

the shading parts are more interesting. You can make better shading. Show me the dark parts.

(Student did so). No look back at your drawing.

All you can see is an outline. Where is the third-dimension?

Turn your page over. I'll hold it and you sketch it. (Teacher held model at three-quarter view).

Sketch - don't just draw a line.

This exchange indicated to the researcher that the teacher had a pre-conception about a single "most



interesting view" for the object. However, later discussion with the teacher revealed that she requested Richard to change his view because children always draw things from the top and never as you "really" see them. Again the emphasis was demonstrated to be upon the teacher's perception of the "facts" (how one "really" sees) rather than upon the meaning of the work for the student.

In light of these kinds of observations the researcher posed the second evaluation question: How well are "non-art" elementary teachers able to direct art appreciation experiences given the open format of the guide materials? The nature of the difference between the interpretation of materials by teacher A and teachers B (reported above) suggested to the researcher that the non-art teachers could be uncomfortable with the non-specificity of outcomes suggested in the guide materials. The use of a factual, linear approach could be indicative of reliance upon the "usual" instructional modes borrowed from other disciplines to compensate for lack of judgemental abilities that are derived from an artistic background.

This hypothesis was subsequently confirmed through discussion with the two Classroom B teachers and the remainder of the people involved with the pilot. Of the 21 pilot teachers, 14 were classified as non-art (possessing less than the equivalent to a minor degree



requirement in art). Nine of that number expressed the opinion that expected outcomes were not stated clearly enough. Each reported that this constituted a frustrating problem for them in that they were unable to judge the effectiveness of lessons and value of outcomes.

Particular difficulties were reported also by these people in questioning students about art works. "What kinds of questions should be asked?" was one concern and "Which are acceptable answers and which are not?" was another.

All of the "art" teachers, on the other hand, expressed the view that outcomes were stated clearly enough, and that non-specificity was important to preserve the flexibility necessary in art education. These teachers also felt confident about judging the outcomes of lessons.

### Presentation of the Findings

The preceding discussion served to illustrate the kinds of evidence sought, procedural methods used and some examples of judgements made. This information was presented to the EPSB in the form of a "qualitative exposition". This required arrangement of judgements and evidence to address each of the evaluation questions presented previously.

Each of the evaluation judgements was supported by evidence, either visual, audio-visual, written



(transcriptions), or interview data. The samples of data presented in the previous section serve to illustrate the nature of the evidence used. Consequently a full presentation of the evaluation report is not required here to illustrate the "qualitative exposition" component of the model. A summary of the most important conclusions, relative to each evaluation question is presented in Appendix"F".

### Generalizability of Findings

Two related questions of generalizability are raised by the EPSB evaluation study. The first concerns representativeness of the classroom experiences observed in two different situations. The second concerns the traditional questions of prediction from a pilot sample to a larger population.

The first issue of representativeness of the two classrooms was created by chance rather than design. (The reasons were explained in Chapter III). In the ideal situation sufficient observers should have been available to view lessons within each of the pilot classrooms. This would have eliminated the need to generalize in the first instance. As well, the study would have been enriched by a variety of judgements relative to the widest possible range of outcomes.

Nevertheless evaluations all suffer some kinds of logistical problems and are seldom conducted in an





"ideal" fashion. The issue of representativeness was therefore addressed in the EPSB study by pursuing issues (identified in the specific classroom situations) through interview with the remainder of the pilot teachers. For example, the issue of "linearity" (discussed previously with respect to non-art teachers' desire for more specific objectives) proved a widespread issue when followed up through interview.

Although the observer's presence in all classrooms, combined with the interview schedule followed in this study would have yielded more information, the procedure actually followed established at least a degree of representativeness.

Another problem of representativeness created through restricting observations to a limited number of classrooms was the effect of limiting the coverage of materials. In other words only a limited sample of the proposed materials was observed in operation.

Again, in the ideal situation, observation in a larger number of classrooms would have eliminated this difficulty. For the purpose of the evaluation however, it was only possible to i) assume consistency of the materials, as the entire set was created by the same writers, and ii) solicit pilot teachers' opinions regarding consistency (pilot teachers agreed unanimously that the materials were largely consistent).

The usual question of prediction should be



considered in light of the nature of the programme content and the methodologies used to derive data for the making of judgements. An assessment of the EPSB materials within the empirical-analytical framework would require pilot testing of the materials in a randomly selected sample of classrooms. Effects of the programme would be measured by means of testing or systematic recording. The probability of these effects occurring in the larger population of classrooms would then be predicted on the basis of the pilot results.

This kind of logic, however, is antithetical to the concept of the programme under investigation. If predictable and consistent results across teachers, students or schools were desired in this art programme (or any other), the nature of those results would be stated in the form of specific behavioural objectives. Instead, vagueness of objectives was deliberately preserved to allow greater individual freedom (of both teachers and students) for interpretation of the meaning of suggested activities and subsequent conduct of artistic inquiry.

The predictive ability of evaluative research in the arts is consequently limited. That the outcomes of a particular programme were systematically observed and judged to be valuable does not necessarily mean that the same specific outcomes are predictable in the future, or in some other place given similar conditions. Nor is it desirable that this should be so. Rather the question



asked of evaluative research in the arts is much more general. Does the programme constitute a fertile set of ideas (stimuli), or framework, for the production of worthwhile and diversified artistic experiences?

In the context of the EPSB project the overarching evaluation question was to assess the potential value of the materials. The content was judged to have been interpreted so that worthwhile activities resulted in the pilot test. In light of the preceding discussion the only predictive statement possible is that the EPSB art appreciation guide materials have the potential for worthwhile interpretation in the wider school system. Further evaluation should occur however, to determine this, following full implementation of the programme.

#### Pragmatic Value of the Model

A brief summary statement is required at this point to reflect upon the pragmatic value of the model. The first question under Pragmatic Norms was "How well does the model function?" The response to this question is favourable in that no major difficulties were encountered in its use. In fact the structure of the model proved particularly useful in two areas. First, no evaluation issues, other than the major one (of determining potential value of the curriculum materials) were specified beforehand by the committee. Second, the nature of the materials themselves was such that specific



outcomes were not delineated. Consequently, the interpretive methodology implicit within the open framework of the model was effectively used to isolate issues and make judgements of value.

It must be emphasised that judgements recorded for this evaluation were the researcher's alone. Although additional judges were not available for classroom observation it has been mentioned that a graduate student was employed to view video -tapes of lessons. This procedure was discontinued after three viewings for the following reasons.

1. No significant differences between the judgements made by the viewer and the evaluator were noted.

2. As the study progressed, "interactive" observation became more significant than "non-interactive". Consequently, judgements made by an external video-tape viewer were limited in terms of the nature of data available for analysis. In other words, the viewer felt difficulty in making judgements, other than general observations, without the benefit of discussion with the teacher and students to determine values and meanings existing in the "live" situation.

3. The video-tape viewer's attention to events was largely pre-determined by the aspect of the camera. Consequently the value of an independent judgement was considerably reduced by eliminating the second observer's freedom of choice. This also was felt to be a factor





contributing towards agreement between observer and evaluator.

Although the proposed model does not rely upon consensus for validity in the empirical-analytical sense, it should be noted that a single set of judgements would be unusual, and not particularly desirable in the normal evaluation situation.

In the event that conflicting judgements be entered by two or more classroom observers, it is proposed that the task of the evaluator would be to choose the judgement most strongly supported by evidence from the situation. It would then be his responsibility to submit the conflicting judgements to decision-makers, together with his own claim and reasons supporting it. This particular procedure is derived from the notion of referential adequacy (a judgement, to be referentially adequate, must be supported by evidence from the work that is apparent to the viewer).

If the decision-makers are unable to accept the evaluator's judgement, the latter must collect further evidence or undertake educative procedures to make clear the reasons supporting his judgements.

Because this practical issue could not be explored in the EPSB study, the answer to the question of "functionality" of the model must be conditional upon further testing.

To the second question under Pragmatic Norms,



"Does the model serve evaluative purposes?" the answer is affirmative if one accepts the definition of evaluation presented earlier in this chapter. The reader will recall that the three essential components of that definition were delineating, obtaining, and presenting useful information to decision-makers to assist consideration of decision alternatives. Earlier in this chapter the delineation component was demonstrated in the form of six evaluation questions, presented under Purposes of the Evaluation. These questions were determined through juxtaposition of the paradigmatic criteria of the evaluator and actors in the social (curriculum) context. Both sets of criteria are evident in the justification of the questions.

The obtaining component was explained at length in Derivation of the Findings. In addition, examples of the evidence used to support judgements were presented in this section. The judgements, and the evidence, arranged in the form of a "qualitative exposition" fulfilled the presenting component of this definition.

The third question under Pragmatic Norms was "What is the nature and usefulness of the information generated by the unique characteristics of the model?"

The nature of the information generated by the model has been fairly thoroughly described elsewhere. In brief the significant types of information gained and their usefulness are:



1. Judgements. The judgemental information provided some indication to the EPSB decision-makers of the value of teacher interpretation and implementation of the curriculum guide materials.

2. Explanations of reasons for action and the meaning of curriculum materials for actors. This information, in addition to assisting with judgements of the value of the programme, also provided information to enable improvement of the materials.

3. Illustrative data of individual behaviours and events. Analytic techniques of the model did not demand subordination of "highlight" or "typical" events and behaviours to categorized (generalizable) data. Consequently these data served particularly useful ends in the presentation of evaluative information.

4. Descriptive and Interpretive information about teaching styles. In an art programme, separation of the teacher and programme materials in the course of examining outcomes is virtually impossible as well as undesirable. The information generated in the EPSB evaluation relative to instructional methods and style, while not immediately applicable in terms of the purposes of the evaluation, demonstrated enormous potential for the improvement of instruction. Limited information which was made available to teachers (video-tapes etc.) was recognized by these individuals as a valuable method of personal analysis.



## Summary of Chapter VI

In this chapter the validity of the model was demonstrated. Both theoretical and empirical evidence was employed to address validation questions raised by the norms of correspondence, coherence, and pragmatism.

The model was presented as two components, a propositional unit (the process of evaluation), and an empirical unit (the sub-model of curriculum development). The sub-model was shown to be empirically derived, and thereby satisfied, to a limited extent, the norm of correspondence.

The model was shown to satisfy the norm of coherence in that it i) demonstrates a similar basic definition of processes of evaluation possessed by other models; ii) is designed to perform similar roles to other models; iii) encompasses within its structure definitions of the components of curriculum contained by other models and iv) contains no obvious contradictions amongst the set of propositions that constitute the model.

The model was shown to possess the unique characteristics of i) availability of an open frame of reference, ii) availability of interpretive methodologies, and iii) availability of analysis techniques that preserved individual and small groups of actions.

Questions raised by pragmatic norms were answered by presenting a brief report and discussion of an





evaluation conducted for the Edmonton Public School Board. In addition to demonstrating the pragmatic value of the model, the EPSB example also served to illustrate the specific nature of model functions. Included in the discussion of this evaluation project were descriptions of i) background to the evaluation project, ii) the evaluation questions derived for the study, iii) methodologies employed both in and out of the classroom, iv) validation concepts appropriate to the methodologies employed, v) design of the evaluation, vi) derivation of the findings, vii) presentation of the findings and viii) problems of generalizability.

The model was found to be useful in the context of the EPSB evaluation. The open framework and interpretive methodologies were found to function particularly well. Evaluative purposes were also served by the model in terms of defining, obtaining and supplying information to decision-makers.

Finally, the nature of information produced by the unique characteristics of the model was found to include i) Judgements, ii) Explanations of reasons for action and the meaning of curriculum materials for actors, iii) Illustrative data of individual behaviours and events iv) Descriptive and interpretive information about teaching styles.



# REFERENCES

- Aoki, T. A Response to Dr. W. James Popham's Paper entitled "Curriculum Design: The Problem of Specifying Intended Learning Outcomes". Presented at a symposium on programme development in education at the University of British Columbia, May 24, 1973.
- Bradley, W. Who Should Evaluate Art Students? Art Education, 1969, 22(9), 22-23.
- Cronbach, L.J. Evaluation for Course Improvement. Teachers College Record, 1963, 64, 676-683.
- Eash, M.J. Developing an Instrument for the Assessment of Instructional Materials (Form IV). Paper presented at the Annual Meeting of the American Educational Research Association, Minneapolis, March 1970. (Worthen, B.R., & Sanders, J.R. Educational Evaluation: Theory and Practice. Worthington, Ohio: Charles A. Jones Publishing Company, 1973)
- Efland, A.D. Evaluating Goals for Art Education. Art Education, 1974, 27(2), 8-10.
- Eisner, E.W. Do Behavioural Objectives and Accountability Have a Place in Art Education? Art Education, 1973, 26(5), 2-5.
- Eisner, E.W. The Perceptive Eye: Toward the Reformation of Educational Evaluation. (Unpublished mimeographed paper of an address presented at a seminar, University of Alberta).
- Kaelin, E.F. Are Behavioural Objectives Consistent with Social Goals of Aesthetic Education. Art Education, 1969, 22(8), 4-11.
- Kaplan, A. The Conduct of Inquiry: Methodology for Behavioral Science. Scranton, Pennsylvania: Chandler Publishing Company, 1964.
- Katahn, M., & Koplin, J.H. Paradigm Clash: Comment on "Some Recent Criticisms of Behaviorism and Learning Theory with Special Reference to Breger and McGaugh and to Chomsky". Psychological Bulletin, 69(2), 147-148.
- Mann, J.S. Curriculum Criticism. Teachers College Record, 1969, 71(1) 27-40.



- Matson, F.W. The Broken Image: Man Science and Society. New York: George Braziller, Inc., 1964.
- Morrisett, I., & Stevens, W.W. Steps in Curriculum Analysis Outline, Boulder, Colo.: Social Science Education Consortium, University of Colorado, 1967. (Mimeo). (Worthen, B.R. & Sanders, J.R. Education: Theory and Practice. Worthington, Ohio: Charles A. Jones Publishing Company, 1973).
- Schatzman, L., & Strauss, A.L. Field Research: Strategies for a Natural Sociology. Englewood Cliffs, New Jersey: Prentice Hall, Inc., 1973.
- Schutz, A. Common Sense and Scientific Interpretation of Human Action. In M. Natanson (Ed.), Alfred Schutz: Collected Papers I: The Problem of Social Reality. The Hague: Martinus Nijhoff, 1973.
- Scriven, M. The Methodology of Evaluation. In R.E. Stake (Ed.), Curriculum Evaluation. American Educational Research Association Monograph Series on Evaluation, 1, Chicago: Rand McNally, 1967.
- Stake, R.E. To Evaluate an Arts Program. In R.E. Stake (Ed.), Evaluating the Arts in Education: A Responsive Approach. Columbus, Ohio: Charles E. Merrill Publishing Company, 1975, 13-32.
- Stewart, W. A Humanistic Study of the Visual Arts, The Commitment, Risk and Potential. Art Education, 1972, 24(2), 2-3.
- Smith, L.M., & Geoffrey, W. The Complexities of an Urban Classroom: An Analysis toward a General Theory of Teaching. Toronto: Holt, Rinehart, and Winston Inc., 1968.
- Sjogren, D.D. Measurement Techniques in Evaluation, Review of Educational Research, 1970, 40, 301-320.



## CHAPTER VII

### Summary, Conclusions and Recommendations

#### Summary

This study represents an attempt to develop and validate a curriculum evaluation model appropriate to the visual arts. Both empirical and theoretical sources of evaluative data and experience were employed to conduct the study.

In Chapter I a lack of attention to evaluation issues was noted in the field of art education. Evaluation models based upon current efficiency concepts of education were viewed to be largely inappropriate for assessment of the visual arts. Research methodologies from alternate paradigms were discussed with respect to their relationship to the epistemology of art. Assumptions of the empirical-analytical research paradigm were found to constitute the major underpinnings of currently used evaluation models, whereas methods of the interpretive sciences were felt by the writer to be more compatible with ways of "knowing" in art.

From the previous discussion emerged the proposition that a curriculum evaluation model appropriate to the visual arts should logically be derived from an artistic (interpretive) paradigm. The unique qualities possessed





by such a model were suggested to be i) availability of an open frame of reference, ii) availability of interpretive methodologies and iii) availability of analysis techniques that do not demand subordination of significant individual outcomes to generalizable information. The concepts of educational connoisseurship and educational criticism, proposed by Eisner (1975), were regarded as useful methodologies upon which to base the model.

Chapter II was comprised of a brief historical review of general evaluation theory, and principles of evaluation in art education. The growth of general evaluation theory through four identifiable stages in this century was noted. The influence on general evaluation theory of developments in psychometrics, the growth of behaviourism, and efficiency concepts of education were described. Popular models of evaluation were also analyzed. The effects of ideas from the field of general evaluation on programme assessment in art education were found to be strong.

In Chapter IV the concept of "validity" was explored at length. Clarification of its meaning was viewed as a critical and necessary endeavour in light of the purposes of the study. Explanation of the concept was required for two purposes. i) To define the norms by which the model developed in this study would be validated and ii) to clarify the particular conceptions of validity employed by research methodologies contributing



to the model. Discussion focussed upon theory validation, and the meaning of validity within the empirical-analytical research framework, and interpretive paradigms of hermeneutics, phenomenology, anthropology, and artistic research.

Chapter V consists of two parts, i) a description of the procedure employed to develop the model and ii) an explanation of its final form. Within the description of procedure the nature of data derived from empirical and theoretical sources was delineated. A pilot test of the model prototype was described and a critical analysis of the results was presented. The rationale which guided development of the final form of the model was outlined.

Explanation of the revised model was conducted in four stages. First, the components of the model were fully delineated. Second, the relationships between the components of the model were explained by means of exploring the analogous relationship of this model for curriculum evaluation with a model for art criticism. Third, the methodological implications of the model were discussed. Finally, conceptions of validity appropriate to the implicit methodologies were outlined.

In Chapter VI the validity of the model was demonstrated. This was achieved by attending to questions raised by the norms of correspondence, coherence and pragmatism. Evidence from empirical and theoretical



sources was employed to satisfy these norms. The major portion of the chapter was devoted to a description of an evaluation conducted for the Edmonton Public School Board. This description served to clarify the functions of the model and to demonstrate its utility.

### Conclusions

Review of the literature revealed that contributions to evaluation theory by art educators have been somewhat sporadic and strongly influenced by theories from other fields. Suggestions regarding appropriate methodologies for evaluation of art programmes were limited and generally non-specific.

In the field of general evaluation theory the development of evaluative techniques was observed to have evolved through four stages. Between 1900 and the 1930's evaluation was equated with administration of standardized tests. From the Thirties to the Fifties evaluation was dominated and typified by the Tylerian tradition. This methodology consisted, in simple terms, of a comparison of outcomes with specifically stated behavioural objectives. During the Sixties the evaluation focus shifted from outcome measurement and comparative evaluation to process assessment. Evaluation became more strongly tied to curriculum development. Programme objectives also became subject to evaluative focus.

In the Seventies some discontent with traditional



concepts of evaluation and curriculum were noted.

"Efficiency" methods of education that employed behavioural approaches were held by some educators to contain serious educational limitations and to confuse efficiency with quality of instruction.

One of the central issues explored in this study was that of "validity". The root meaning of the term was found to be "strength", and valid knowledge has come to be regarded as that which provides the most powerful means to given epistemological ends.

The fugitive nature of the term was revealed by examining its meaning within alternate research paradigms. Its meaning was observed to change according to assumptions of reality adopted by particular schools of philosophy.

Within the empirical-analytical framework, validity of knowledge was found to be based upon the assumption of a world "out there" separate and distinct from the observer. Highly sophisticated rules have been developed to ensure "objectivity" of measurements and observations. These rules are generally based upon procedures of consensus.

The interpretive sciences were observed to assume different positions with respect to concepts of reality. Interpretive methodologies accept interaction of subject and object. Reality is therefore regarded to be socially constructed and not separate from the observer.

The hermeneutic tradition is concerned with





interpretation of text and text-analogues. Three different positions were recorded with respect to criteria held for determination of validity. They were correspondence with the author's intention, construction of intersubjectively shared meanings (sense making) and the discovery of new meanings.

Within the phenomenological paradigm it was noted that the unique criterion for validity was not consensus. Rather, explanations of human actions should be understandable (and recognized as correct interpretations) by the actor and his fellow men in the common sense world of everyday life.

Within the artistic paradigm, interpretation of the meaning of art objects was found to be complicated because the concept of "art" is open ended, artistic knowledge is particular, and art does not attempt to approximate something else (truth or reality) which can be used as a criterion for validity.

Three basic criteria for validity of artistic judgements were observed to be held by aestheticians. They were correspondence with the intention of the artist, correspondence with the meaning implicit within the work, and referential adequacy. The latter was judged to be useful as a means of determining strength of observations made about artistic phenomena.

These conclusions carried very definite implications for construction of this and by implication, similar models.



The most important consideration was viewed to be preservation of interpretive methodologies which recognize "multiple realities" and "socially constructed meanings" of actors in the curriculum context. Validation procedures held to contain the greatest potential for an "art" evaluation model were satisfaction of the "postulate of adequacy" (phenomenology) and "referential adequacy" (art criticism).

Also discussed in Chapter IV in relation to validity were three norms for theory validation. They were the norms of correspondence, coherence, and pragmatism. The norm of correspondence requires a theory to demonstrate empirical connection so that it may be viewed to "fit the facts". The norms of coherence and pragmatism are generally subordinate to this notion. Satisfaction of the norm of coherence demands that new theories integrate with those already existing, and possess internal logic, and an appropriate degree of complexity. The pragmatic norm is simply a requirement of practical value (scientific utility) and is not in itself a necessary or sufficient condition for validation. These norms were considered to be significant in terms of their usefulness for demonstrating validity of the model.

In Chapter V the general principles referred to before as necessary for a model of curriculum evaluation in the visual arts were developed in more specific terms. Analysis of two curriculum projects, review of the literature, and a critical analysis of the results of a



pilot test of the prototype of the model, produced a set of ideas which were fashioned into the final form of the model.

Study of the curriculum projects and the literature permitted identification of six curriculum elements, i) agents, ii) content, iii) implementation, iv) outcomes, v) rationale, and vi) resources.

From general evaluation theory three central ideas were identified. They were: i) evaluation is a process of defining, obtaining, and providing useful information to decision-makers, ii) evaluation is a process of social valuing, iii) evaluation should play a role in curriculum improvement. Three further ideas were isolated from the field of art education. They were: i) specific objectives need not be stated in advance to ensure quality art programmes, ii) lesson content in art programmes does not have to be pursued in a pre-determined fashion to be worthwhile, iii) desired student responses in art are individual rather than normative.

In addition to the unique characteristics required of the model (open frame of reference, interpretive methodologies, and analysis techniques which preserve the individual event) three other features were identified from the literature, and the pilot test, as important considerations for model construction. They were i) flexibility of viewpoint, ii) flexibility of role and iii) flexibility of methodology.



Revision of the model following pilot testing of the prototype was necessary to eliminate implied linearity and inappropriate conceptions of logical contingency and congruence. The prototype was found also to lack the desired "open framework", and inadequately accommodated this social dimension of the evaluation process.

Prior to the revision, similarities were noted between curriculum evaluation and art criticism. These similarities were: i) both critic and evaluator perform value judgements for a specific audience, ii) both curricula and art objects depend for their structure and content upon numerous decisions made by their creators, iii) the forms of both phenomena are characterized by a discrete number of generic elements, iv) both art and curricula are open concepts.

These similarities led to the proposition of a model for curriculum evaluation that displayed similar relationships to a model for art criticism. The previously identified elements of curriculum were preserved in the revised model. The fundamental evaluative relationship was conceived to result from juxtaposition of two sets of paradigmatic criteria, the evaluator's and the actor's (curriculum developers, administrators, teachers, and students). The product of the evaluative process was called "qualitative exposition". The revised model was claimed to accommodate the ideas defined previously and correct the deficiencies displayed by the prototype.

The most appropriate methodologies implied by the





model were viewed to be interpretive, although the possibility that quantitative analysis may be demanded by different types of programmes was not ignored. The "paradigmatic criteria" of the evaluator were proposed to serve in selection of evaluation foci and determination of the worth of outcomes.

The essential difference between criticism of an art work and evaluation of a curriculum was explained in terms of the difference between a static object which contains no inherent meaning and a dynamic social phenomenon which is comprised of actors who create meaning within their social world.

The types of data available to an evaluator for collection were identified as subjective and objective. Subjective data were defined as actors' explanations of their reasons for action and classroom critics' judgements of the value of observed events. Objective data were defined as external information about curriculum events such as that recorded by mechanical devices or observers operating in the empirical mode. Because of the difference noted between art and curriculum it was proposed in this model that appropriate judgements made by classroom observers should be made on the basis of both external and subjective data.

Analysis of data was observed to be possible in terms of "causes" or "reasons". It was argued that analysis in terms of "reasons" was the approach most compatible with the epistemology of art as it accepts free will of the



individual (a basic tenet of artistic behaviour).

Three useful conceptions of validity were viewed to exist in connection with the data types outlined above. Objective data that are generated by observers operating in the empirical mode rely upon consensus techniques for validity. Judgemental data produced by classroom observers must demonstrate referential adequacy, and phenomenological data (actors' explanation of the meaning of their actions) must satisfy the postulate of adequacy.

Validation of the model was undertaken in Chapter VI. Validity was demonstrated by attending to questions raised by the norms of correspondence, coherence and pragmatic norms.

The basic question asked by the norm of correspondence was "Does the theory fit the facts?" In response to that question it was argued that the model was comprised of empirical and postulational components, and, as the postulational component represented a proposition for evaluation there seemed no point in attempting to validate it on empirical grounds. Within the total framework however, the curriculum development dimension was demonstrated to be derived empirically. In response to the question of "fitting the facts" it was suggested that at least the sub model possesses a certain degree of empirical validity and therefore fits the norm of correspondence.

Three questions were dealt with under the norm of coherence. The first was "How does the model integrate



with existing theories of curriculum evaluation?" Following examination of the literature it was concluded that the model could claim some integration with existing evaluation theory because i) it demonstrates a similar basic definition of process of evaluation possessed by other models, ii) it is designed to perform similar roles to other models (formative and summative), iii) it encompasses with its structure definitions of the components of curriculum implied by other models.

The second question under the norm of coherence was "In what respects does the model differ from existing models?" The distinguishing characteristics of the model, stated as necessary requirements in Chapter I, represent the ways in which it differs from others. These characteristics were listed before as i) availability of an open frame of reference, ii) availability of interpretive and judgemental methodologies, iii) availability of analysis techniques that preserve individual or small groups of behaviours.

The meaning of these characteristics was explained as follows:

1. The open frame of reference means that neither evaluative nor curriculum development sequences are pre-determined. Rather, the open framework accepts that the means and ends of art education can, and usually do, exist in an interactional relationship. They are not clearly separable as in the Tylerian mode. The major



implication of this is that a clear picture of the end product (i.e. individuals who are educated in the arts) is not available to evaluators or curriculum developers.

2. Use of judgemental and interpretive methodologies means that traditional conceptions of validity do not apply in the same manner to data gathering and analysis methods. The kinds of validation concepts that are applicable were discussed previously.

3. Preservation of individual or small groups of behaviours or events in analysis of the data is a desirable characteristic in that it allows the evaluator to highlight significant points in the "qualitative exposition". The "qualitative exposition" component is in itself a unique procedure which reflects the salient characteristics of art criticism.

The third question asked by the norm of coherence was: "Is the logic of the model internally consistent?" The only satisfactory response which could be made to this question was that there are no obvious contradictions operating amongst the set of propositions which constitute the model. No other response could be made, given the sense in which "model" is used in the study (i.e. as a set of propositions for actions rather than as an explanation of phenomena).

Prior to answering the questions posed by the pragmatic norms a description of an evaluation conducted for the Edmonton Public School Board was undertaken. This





description served to illustrate the functions of model components as well as to demonstrate its utility.

Under the question "How well does the model function?" two observations were made.

1. No difficulties were encountered in its use. A cautionary note should be entered here in that the evaluation project itself contributed towards the form of the model. To test the model in a situation which contributed in any way to its development could have contributed to a certain degree of circularity. However, it is argued that sufficient information from the FAC programme and from the literature was employed to minimise this possibility.

2. The open framework of the model provided sufficient flexibility for the evaluator to effectively isolate evaluative issues and make judgements about programme outcomes. However, because the evaluation was conducted largely by the researcher the possibility of conflicts of judgements was not able to be dealt with. The significance of the researcher's own art background was found to play a large part in determining evaluation foci both in and out of the classroom. The suggestion is made that non-art personnel would be unable to act as classroom observers or evaluators due to their lack of useful paradigmatic criteria. This observation closely parallels the issue that arose in the EPSB study with respect to non-art teachers' use of "open ended" materials. Because of their



lack of art background these teachers expressed the view that they were unable to determine which questions to ask in discussion, or how to judge the outcomes of the programme.

To the second question under pragmatic norms "Does the model serve evaluative purposes?", the answer was affirmative. Procedures suggested by the model served to delineate, obtain, and present evaluative information to decision-makers.

The third question under pragmatic norms was: "What is the nature and usefulness of the information generated by the unique characteristics of the model?"

Significant types of useful information provided by the unique characteristics were i) judgemental information regarding the value and usefulness of teacher interpretation and implementation of curriculum guide materials, ii) explanations of reasons for actions and the meaning of curriculum materials for teachers and students iii) illustrative data on individual behaviours and events, iv) descriptive and interpretive information about teaching styles.

One possible limitation of the model was observed to exist in relation to generalizability. Judgements of the value of observed events do not possess predictive power in the tradition of empirical - analytical research. That is, specific programmes may be judged to produce worthwhile results in a sample situation but prediction



of the specific nature of outcomes in a larger population is not possible or even particularly desirable. For this reason it is suggested that the model is probably more appropriate to smaller projects where larger proportions of the outcomes may be viewed. On the other hand the case-study approach has been used previously in education to produce useful particular information for a more complete understanding of specific types of educational endeavours. In any event generalizing from a sample to the population is a questionable activity in light of the changes which inevitable occur in the conditions of education. In addition, the elusive nature of the true sample is a condition which has traditionally plagued empiricists in the social sciences.

To demonstrate validity of the model beyond all doubt is a task beyond the scope of any single researcher, or group of researchers for that matter. Validity is therefore claimed subject to further testing, use and modification by others in the field.

### Recommendations

That the model should be tested in a variety of situations is the most obvious implication of the study. The EPSB project was particularly suitable for the initial test of the model. The nature of that programme was such that use of traditional models would have been inappropriate for the reasons outlined in Discussion of the Problem in



## Chapter I.

Testing of the model in situations requiring a variety of classroom critics is a necessary next step. The difficulty of reconciling possible conflicting judgements is an intriguing question which requires extensive explanation. It was recommended in the body of the study that both evaluator and classroom critics should possess appropriately extensive background in the visual arts (both in education and studio). This is the essential requirement for practical implementation of the model. Without this background the "paradigmatic criteria" component of the model becomes ineffective. The classroom critics require art knowledge (in the tradition of the interpretive sciences) to make judgements about the value of observed events. The evaluator must have extensive art knowledge to weigh the relative strength of his observer's judgements, and to assist in the refinement of his clients' perceptions where necessary.

The difficulty of limited predictive power, inherent within artistic research, suggests that the model is most appropriate for assessment of smaller projects. An enormous amount of judgemental data related to teaching style was produced during the course of the EPSB project which could not be used for the purposes of that particular evaluation. However, those data contained enormous potential for use in improving the educational experiences within particular classrooms.





The indefinite nature of most art curricula, while ostensibly intended to allow flexibility of choice for teachers, serves to place ultimate responsibility for the quality and appropriateness of programme content and method upon the teacher. Art programmes generally become teacher, or teacher/student designed sequences of exercises. Selection of these exercises is critically dependent upon the teacher's evaluative skill. Working within the broad guidelines that constitute the curriculum guide, the teacher is constantly responsible for assessing the quality of student's critical and productive performance. The products of his evaluation then serve to assist in the making of decisions about future activities for individuals or groups of students. Variation in teacher ability to make evaluative judgements will be reflected in the appropriateness of student experiences.

The implication from this argument is that particular educational benefits could be derived through use of the model in small scale situations (single schools or groups of schools) to assist non-art teachers to deal with the flexible nature of most art curricula.

Despite the fact that generalizability of judgemental data is limited, the model is not necessarily restricted to small scale evaluations. The possibility of utilizing internal critics (school system experts) for classroom judgemental activity could be explored. The advantage of this course of action would be in the



possibility of wide coverage of observations.

The model developed within this study represents a unique contribution to evaluation theory in art education. For the first time an attempt has been made to draw together a set of ideas from art education and general evaluation theory to create a distinctive type of curriculum evaluation model appropriate to the visual arts. It may be used in place of, or perhaps as a complementary form to other methods of evaluation. Its strength lies in the fact that its methods are compatible with the epistemology of art. Because the model allows examination of educational activities which proceed in a purposive, rather than a purposeful manner, it may have application to other subject areas of the school curriculum possessing similar characteristics.



B I B L I O G R A P H Y



- AAAS Commission on Science Education. An Evaluation Model and its Application. Washington D.C.: American Association for the Advancement of Science. Publication 65-9, 1965.
- Alkin, M.C. Evaluation Theory Development. Evaluation Comment, 1969, 2(1), 2-7.
- Aoki, T. A Response to Dr. W. James Popham's Paper entitled "Curriculum Design: The Problem of Specifying Intended Learning Outcomes". Presented at a symposium on programme development in education at the University of British Columbia, May 24, 1973.
- Apple, M.W. The Process and Ideology of Valuing in Educational Settings. In M.W. Apple, M.J. Subkoviak, & H.S. Lufler Jr., (Eds.), Educational Evaluation: Analysis and Responsibility. Berkely: McCutchan, 1974.
- Atkin, J.M. Some Evaluation Problems in a Course Content Improvement Project. Journal of Research in Science Teaching, 1963, 1, 129-132.
- Beardsley, M.C. Aesthetics: Problems in the Philosophy of Criticism. New York: Harcourt, Brace & World Inc., 1958.
- Beardsley, M.C. The Classification of Critical Reasons. In R.A. Smith (Ed.) Aesthetics and Problems of Education. Chicago: University of Illinois Press, 1971.
- Beittel, K.R. Art. In Chester W. Harris, (Ed.), Encyclopedia of Educational Research. New York: The Macmillan Company, 1960.
- Beittel, K.R. Formative Hermeneutics in the Artting Processes of an Other: The Philetics of Art Education, Art Education, 1974, 27(9), 2-7.
- Berger, P.L., & Luckmann, T. The Social Construction of Reality: A Treatise in the Sociology of Knowledge. New York: Doubleday & Company Inc., 1966.
- Bloom, B.S. Taxonomy of Educational Objectives: Cognitive Domain. New York: David McKay Co., Inc., 1956.
- Bobbitt, F. How to Make a Curriculum. Boston: Houghton Mifflin, 1924.
- Bonner, H. Two Views of Science. On Being Mindful of Man: Essay Toward a Proactive Psychology. Boston: Houghton Mifflin Co., 1965.





- Bradley, W. Who Should Evaluate Art Students? Art Education, 1969, 22(9), 22-23.
- Campbell, D.T., & Stanley, J.C. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally College Publishing Company, 1963.
- Cattell, R.B. Validity and Reliability: A Proposed More Basic Set of Concepts. Journal of Educational Psychology, 1964, 55(1), 1-22.
- Child, I.L. The Problem of Objectivity in Esthetic Value. PennState Papers in Art Education, 1966, 1.
- Cicourel, A.V. Cognitive Sociology: Language and Meaning in Social Interaction. London: Penguin Books Ltd., 1973.
- Conrad, G. The Process of Art Education in the Elementary School. Inglewood Cliffs, N.J.: Prentice Hall Inc., 1964.
- Cronbach, L.J. Evaluation for Course Improvement. Teachers College Record, 1963, 64, 676-683.
- Cureton, E.E. Validity Reliability and Baloney. Educational and Psychological Measurement, 1950, 10, 94-96.
- D'Amico, V. Millennium or Mirage. Art Education, 1966, 19(5), 27-32.
- Davis, D.J. Human Behaviour: Its Implications Curriculum Development in Art. Studies in Art Education, 1971, 12(3), 10-12.
- Day, M. Rationales for Art Education: Thinking Through and Telling Why. Art Education, 1972, 25(2), 17-20.
- Dearden, R.F. The Aesthetic Form of Understanding. In R.A. Smith (Ed.), Aesthetics and Problems of Education. Chicago: University of Illinois Press, 1971.
- De Francesco, I.L. Art Education, Its Means and Ends. New York: Harper Brothers Publishers, 1958.
- Dewey, J. Art as Experience. New York: G.P. Putnam's Sons, 1934.
- Dimondstein, G. Exploring the Arts with Children. New York: McMillan Publishing Co. Inc., 1974.
- Dobbs, S.M. Program Analysis in Arts Education. Studies in Art Education, 1972, 12(2).



- Eash, M.J. Developing an Instrument for the Assessment of Instructional Materials Form IV. Paper presented at the annual meeting of the American Educational Research Association, Minneapolis, March, 1970. (Worthen, B.R., & Sanders, J.R. Educational Evaluation: Theory and Practice. Worthington, Ohio: Charles A. Jones Publishing Company, 1973).
- Ecker, D.W., & Kaelin, E.F. The Limits of Aesthetic Inquiry: A Guide to Educational Research. In P.L.G. Thomas (Ed.), Philosophical Redirection of Educational Research. The Seventy-First Yearbook of the National Society for the Study of Education. Part I. Chicago, Illinois: University of Chicago Press, 1972.
- Efland, A.D. Evaluating Goals for Art Education. Art Education, 1974, 27(2), 8-10.
- Eisner, E.W. Educating Artistic Vision. New York: The Macmillan Company, 1972.
- Eisner, E.W. Do Behavioural Objectives and Accountability Have a Place in Art Education? Art Education, 1973, 26(5), 2-5.
- Eisner, E.W. Research on Teaching the Visual Arts. In M.W. Travers (Ed.), Second Handbook of Research on Teaching. Chicago: Rand McNally College Publishing Co., 1973, 1196-1209.
- Eisner, E.W. Toward a More Adequate Conception of Evaluation in the Arts. Art Education, 1974, 27(7), 3-5.
- Eisner, E.W. Some Alternatives to Quantitative Forms of Educational Evaluation. Thrust, 1975, 5(2), 13-15.
- Eisner, E.W. The Perceptive Eye: Toward the Reformation of Educational Evaluation. (Unpublished mimeographed paper of an address presented at a seminar, University of Alberta), 1975.
- Feldman, E.B. Becoming Human through Art. Englewood Cliffs, New Jersey: Prentice Hall, Inc., 1970.
- Habermas, J. Knowledge and Human Interests. Boston: Beacon Press, 1971.
- Hammond, R. Context Evaluation of Instruction in Local School Districts. Educational Technology, 1969, 9(1), 13-18.
- Handley, W.B. Substantive, Structural and External Components of Construct Validity in Instrumentation. Unpublished doctoral dissertation, University of Alberta, 1973.



- Harris, I.M. Agreement in Art: An Educational Issue: Journal of Aesthetic Education, 7(3), 1973, 51-61.
- Hausman, J.J. Research on Teaching the Visual Arts. In N.G. Gage (Ed.), Handbook of Research on Teaching, Chicago: Rand McNally, 1963, 1101-1117.
- Hausman, J.J. Curriculum Projects: Reflections on their Evaluation and Continuance. Art Education, 1971, 24(8), 24-27.
- Hirsch, E.D., Jr. Validity in Interpretation. New Haven: Yale University Press, 1967.
- Hubbard, G. Art in the High School. Belmont, California: Wadsworth Publishing Co. Inc., 1967.
- Kaelin, E.F. Are Behavioural Objectives Consistent with Social Goals of Aesthetic Education. Art Education, 1969, 22(8), 4-11.
- Kaplan, A. The Conduct of Inquiry: Methodology for Behavioral Science. Scranton, Pennsylvania: Chandler Publishing Company, 1964.
- Karier, C.J. Ideology and Evaluation: In Quest of Meritocracy. In M.W. Apple, M.J. Subkoviak, & H.S. Lufner Jr., (Eds.), Educational Evaluation: Analysis and Responsibility. Berkeley: McCutchan, 1974, 279-319.
- Katahn, M., & Koplin, J.H. Paradigm Clash: Comment on "Some Recent Criticisms of Behaviorism and Learning Theory with Special Reference to Breger and McGaugh and to Chomsky". Psychological Bulletin, 69(2), 147-148.
- Knudsen, E.H., & Christensen, E.M. Children's Art Education. Peoria, Illinois: Chas. A. Bennett Co., Inc., 1957.
- Krathwohl, D.R., Bloom, B.S., & Masia, B.B. Taxonomy of Educational Objectives: Affective Domain. New York: David McKay Co., Inc., 1964.
- Kuhn, T.S. The Structure of Scientific Revolution. Chicago: Chicago:University Press, 1970.
- Kuhns, R. Criticism and the Problem of Intention. In R.A. Smith(Ed.), Aesthetics and Criticism in Art Education. Chicago: Rand McNally & Company, 1966.
- Lansing, K.M. Art, Artists and Art Education. New York: McGraw-Hill, 1971.
- Larson, C.H. Search and Research. Art Education, 1966, 19(3).
- Lewis, H.P. Point. Studies in Art Education, 1972, 13(3).





- Loevinger, J. Objective Tests as Instruments of Psychological Theory. Psychological Reports, 1957, 3, 635-694.
- Lowenfeld, V. Creative and Mental Growth. New York: The Macmillan Company, 1947.
- Mager, R.F. Preparing Instructional Objectives. Palo Alto, California: Fearon Publishers, 1962.
- Mann, J.S. Curriculum Criticism. Teachers College Record, 1969, 71(1), 27-40.
- Matson, F.W. The Broken Image: Man Science and Society. New York: George Braziller, Inc., 1964.
- Messick, S. The Standard Problem: Meaning and Values in Measurement and Evaluation. Presidential Address to the Division of Evaluation and Measurement of the American Psychological Association at New Orleans, August 30, 1974.
- Morrisett, I., & Stevens, W.W. Steps in Curriculum Analysis Outline, Boulder, Colo.: Social Science Education Consortium, University of Colorado, 1967. (Mimeo). Worthen, B.R., & Sanders, J.R. Education: Theory and Practice. Worthington, Ohio: Charles A. Jones Publishing Company, 1973).
- Natanson, M. Introduction. In M. Natanson (Ed.), Alfred Schutz: Collected Papers I: The Problem of Social Reality. The Hague: Martinus Nijhoff, 1973.
- Pohland, P. Participant Observation as a Research Methodology. Studies in Art Education, 1972, 13(3).
- Polanyi, M. The Logic of Liberty. Chicago: University of Chicago Press, 1951.
- Polanyi, M. The Study of Man. Chicago: University of Chicago Press, 1959.
- Polanyi, M. Personal Knowledge. New York: Harper and Row, 1964.
- Popham, W.J. Curriculum Design: The Problem of Specifying Intended Learning Outcomes. Presented at a Symposium on programme development in education at the University of British Columbia, May 24, 1973.
- Provus, M. Evaluation of Ongoing Programs in the Public School Systems. In R.W. Tyler (Ed.), The Sixty-Eighth Yearbook of the National Society for the Study of Education, Part II, Chicago: The University of Chicago Press, 1969, 242-83.





- Schatzman, L., & Strauss, A.L. Field Research: Strategies for a Natural Sociology. Englewood Cliffs, New Jersey: Prentice Hall, Inc., 1973.
- Schutz, A. Common Sense and Scientific Interpretation of Human Action. In M. Natanson (Ed.), Alfred Schutz: Collected Papers I: The Problem of Social Reality. The Hague: Martinus Nijhoff, 1973.
- Scriven, M. The Methodology of Evaluation. In R.E. Stake (Ed.), Curriculum Evaluation. American Educational Research Association Monograph Series on Evaluation, 1, Chicago: Rand McNally, 1967.
- Scriven, M. Prose and Cons about Goal-Free Evaluation. Journal of Educational Evaluation, 1972, 3(4), 1-4.
- Smith, E.R., & Tyler, R.W. Appraising and Recording Student Progress. New York: Harper and Row, 1942.
- Smith, E.R., Tyler, R.W., & the Evaluation Staff. The Evaluation and Appreciation of Art. Adventure in American Education 3, New York: Harper & Brothers, 1942.
- Smith, L.M. The Use of Observational Strategies for the Development of a Summative Evaluation Model. Washington University and CEMREL, Inc. Paper presented as part of the Symposium Evaluation Efforts in Laboratory and R & D Centres: Some Emerging Issues. AERA, Chicago, April 7, 1972. (ERIC Document Reproduction Service No. ED064367).
- Smith, L.M., & Geoffrey, W. The Complexities of an Urban Classroom: An Analysis toward a General Theory of Teaching. Toronto: Holt, Rinehart, and Winston Inc., 1968.
- Stake, R.E. The Countenance of Educational Evaluation. Teachers College Record, 1968, 68, 523-540.
- Stake, R.E. To Evaluate an Arts Program. In R.E. Stake (Ed.), Evaluating the Arts in Education: A Responsive Approach. Columbus, Ohio: Charles E. Merrill Publishing Company, 1975, 13-32.
- Steele, C., & Victoria, J. Accountability and Assessment of Art: Education in Public Schools. Art Education, 1975, 28(5).
- Stewart, W. A Humanistic Study of the Visual Arts, The Commitment, Risk and Potential. Art Education, 1972, 24(2), 2-3.



- Stewart, W. Point. Studies in Art Education. 1972, 13(3).
- Stufflebeam, D.L. The Use and Abuse of Evaluation in  
Title III. Theory into Practice, 1967, 6(3), 126-33.
- Stufflebeam, Foley, Gephart, Guba, Hammond, Merriman, &  
Provus. Educational Evaluation and Decision Making.  
Itasca, Illinois: F.E. Peacock Publishers, Inc.,  
1971.
- Sjogren, D.D. Measurement Techniques in Evaluation,  
Review of Educational Research, 1970, 40, 301-320.
- Taba, H., & Sawin, E.I. A Proposed Model in Evaluation.  
Education Leadership, 1962, 20, 57-71.
- Taylor, P.A., & Maguire, T.A. A Theoretical Evaluation  
Model. Manitoba Journal of Educational Research,  
1966, 1, 12-17.
- van Manen, M. Some Search Orientations for Research in  
Social Studies Education. Paper presented at the  
NCSS Research Advisory Committee Session, Chicago,  
1974.
- Weitz, M. The Nature of Art. In D.W. Ecker & E.W. Eisner  
(Eds.), Readings in Art Education, Waltham, Mass.:  
Blaisdell, 1966, 49-56.
- Willis, G.H. Curriculum Criticism and Literary Criticism.  
Journal of Curriculum Studies, 1975, 7(1), 3-17.
- Wilson, B.G. A Proposal for Systematic Curriculum  
Evaluation in Art Education. Studies in Art  
Education, 1968, 9(3).
- Wilson, B.G. Evaluation of Learning in Art Education. In  
B.S. Bloom, J.T. Hastings, & G.F. Madaus (Eds.),  
Handbook of Formative and Summative Evaluation of  
Student Learning. New York: McGraw Hill Publishing  
Company, 1972, 499-558.
- Wilson B.G. Point. Studies in Art Education, 1972, 13(3).
- Wilson, B.G. The Other Side of Evaluation of Art Education.  
In G.W. Hardiman & T. Zernich (Eds.), Curricular  
Considerations for Visual Arts Education:  
Rationale Development and Evaluation. Champaign,  
Illinois: Stipes Publishing Company, 1974, 247-76.
- Worthen, B.R., & Sanders, J.R. Educational Evaluation:  
Theory and Practice. Worthington, Ohio: Charles  
A. Jones Publishing Company, 1973.



Weitz, M. The Nature of Art. In D.W. Ecker & E.W. Eisner  
(Eds.), Readings in Art Education, Waltham, Mass.:  
Blaisdell, 1966, 49-56.

Zuniga, J.A. Critical Analysis and its Subject Matter.  
Journal of Aesthetics and Art Criticism, 1973,  
32(2).



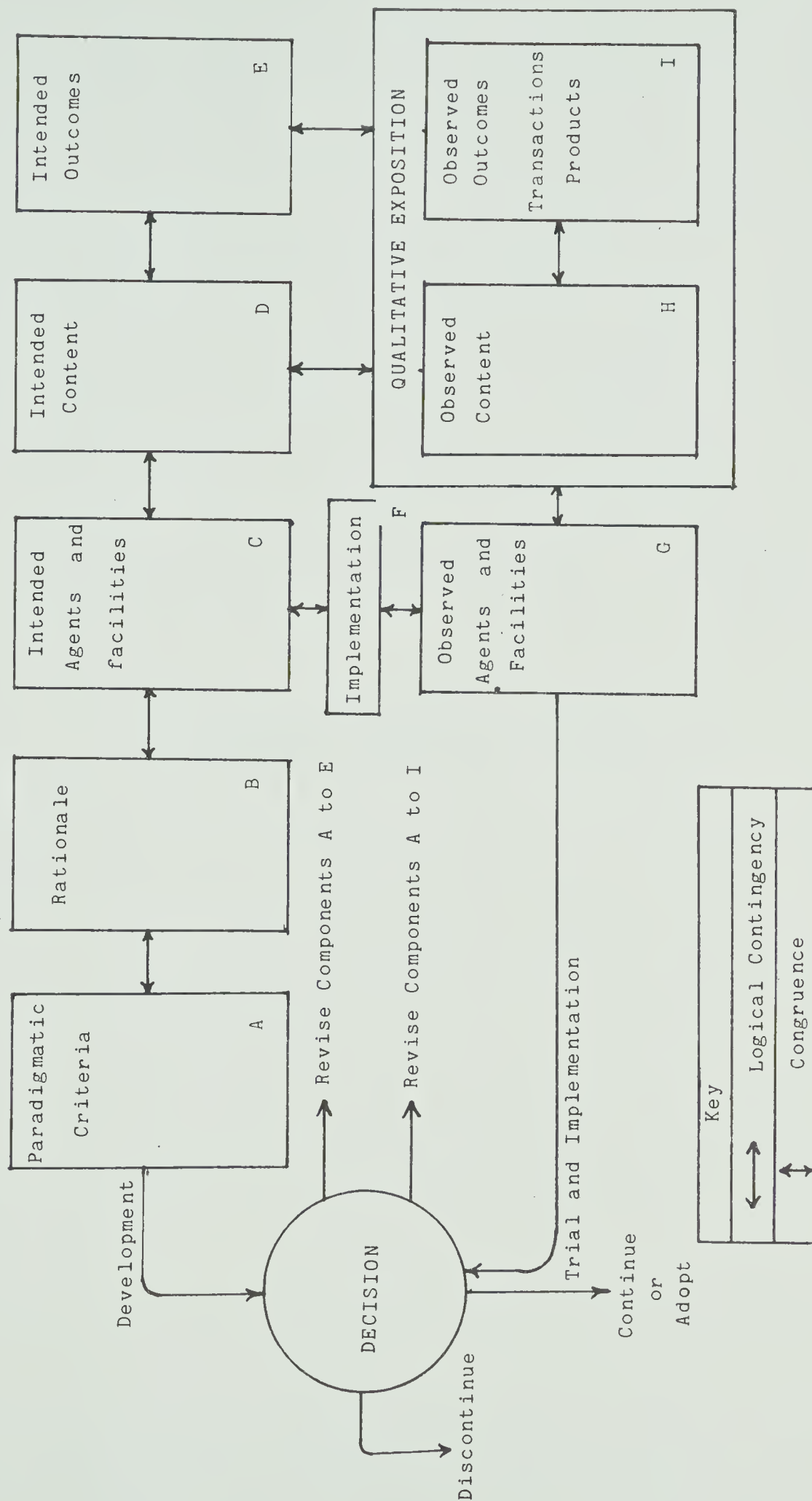
APPENDIX "A"





## Curriculum Evaluation Model for the

## Visual Arts - (Prototype)





APPENDIX "B"



EXPLANATION OF EVALUATION QUESTIONS:  
THE 1975 FINE ARTS COUNCIL CONFERENCE

Question 1 - What needs did the conference serve?

It was assumed an analysis of the professional designation of participants, attracted by the conference agenda, would reflect, to a limited extent, the assumptions made by conference planners with respect to needs served by the conference. As well, this information provided a basis for inference regarding the implicit conference "rationale".

Question 2 - What were the participants' perceptions regarding the value of the conference?

The participants' opinions regarding "overall value" of the conference were particularly significant in the context of this evaluation. The conference is planned essentially by selected Fine Arts Council members for the professional enrichment of the membership at large. Conference planners are therefore charged with the responsibility of building a program that is "valuable" to the participants. This is a key evaluative question in the sense that the worth of the conference is determined, primarily, by the degree to which participants perceive it as valuable.

Question 3 - What are the potential effects of the conference sessions upon participants'



subsequent professional activities?

As the Fine Arts Council is an association operating within the broader context of the ATA, it is reasonable to assume that an implicit intention of a FAC conference would be to achieve the objectives stated in the constitution of the Fine Arts Council. The conference, therefore, if successful, is likely to produce some effect upon the professional activity of the participants. This assumption was supported indirectly by the statement of conference purposes defined by the conference planning committee, as well as the statement of objectives of the FAC in the constitution.

Question 4 - What recommendations for change were made by participants?

The purpose of this question is to allow exploration of specific needs of participants not served by the conference. This information is directly related to Questions 1 and 2 but is rather more specific in nature.

Question 5 - What were the instructors' perceptions regarding interest generated by specific sessions?

Instructors' perceptions regarding degree of interest generated by their sessions could provide useful specific information with respect to the nature of the sessions as well as other variables operating in individual situations. Comprehensive reporting from each instructor





was not anticipated under this question due to the difficulties imposed by the short term nature of the conference, and the shortage of interviewers.

Question 6 - What degree of participation was observed at specific sessions?

The purpose of this question is identical to Question 5. The difference between the two is one of viewpoint (i.e. participant observer versus instructor). Discrepancies between the viewpoints were not necessarily sought.

Rather, a broader range of descriptive information and opinion was desired. The same limitations for data collection existed here as in the previous question. Because of the shortage of observers, all sessions were not observed. Sessions observed were selected according to the interests of the observers.

Question 7 - How effectively were conference plans implemented?

Question 8 - Was the content of specific sessions congruent with participants' expectations?

This question relates specifically to the broader scope of Question 1. Also under this question some exploration of the effects of "pre-session advertising" was undertaken.

Question 9 - Did the conference achieve its purposes?

Conference purposes were defined by the planning committee.



The degree to which they were achieved reflected the appropriateness of the committee's decisions.

Question 10 - How valuable was the conference?

This is the omnibus evaluation question that draws together implications from data generated by the above questions.

Also, prior questions relating to the constitutional function of the FAC within the ATA and objectives for the conference held by the ATA were examined.



APPENDIX "C"



## CONCLUSIONS AND RECOMMENDATIONS:

## FINE ARTS COUNCIL CONFERENCE REPORT

General Conclusions

The conference was perceived as valuable by the participants, planners and the researcher. Some factors contributed towards obfuscation of the value of the event from the evaluator's perspective. Following is a summary of the conclusions for each evaluation question. (See Appendix "B")

1 - (a) The Conference served a broad spectrum of needs.

(b) Art and music still experience a greater demand than drama despite a 50 percent attendance growth in drama since 1973.

(c) Common needs expressed across categories of attenders were:

(i) to gain insight and ideas regarding content in specific areas.

(ii) to explore philosophic issues.

(iii) to engage in social exchange.

(d) A strong impetus to attend was provided the participants by the appeal of the site.

2 - The conference was generally perceived as valuable by the participants.

3 - The conference sessions were perceived as likely to produce some subsequent effect upon professional activity. The nature of this effect was unspecified.

4 - Some recommendations for change were made by participants.





These recommendations were low in frequency in most instances. The strongest recommendations related to:

Publicity

- (a) Publicity could have been more extensive.
- (b) Programs should have been distributed beforehand.

Sessions

- (a) Sessions should be repeated to increase flexibility of choice.
- (b) Sessions should be programmed Friday evening and Sunday morning.

Other recommendations related to: organization, displays and venue.

- 5 - Instructors perceived a high level of interest and responsiveness of participants at sessions.
- 6 - The degree of participation at the sessions was generally high, with the exception of the general sessions where attendance and participation were low.
- 7 - Conference plans were effectively implemented.
- 8 - Content of specific sessions was congruent with participants' expectations in most cases.
- 9 - (a) The conference was highly successful in providing stimulating ideas in art, music and drama.

(b) The conference was successful in:

- (i) providing opportunity for participants to meet others in the province with similar professional interests.
- (ii) providing assistance with individual and



common problems in formal sessions.

(iii) promoting discussion on professional issues in formal small group sessions and informal social sessions.

(iv) allowing participants to acquaint themselves with currently available material resources in music and art.

(c) The conference was less successful in:

(i) making plans for group action to be taken later.

(ii) providing opportunity for discussion across discipline.

(iii) providing assistance with problems in social sessions.

(iv) promoting discussion on professional issues in the general sessions.

(d) The conference was unsuccessful in allowing participants to acquaint themselves with currently available material resources in drama.

10- (a) From the perspective of the values held by participants and planners the conference was valuable.

(b) From the perspective of the evaluator the conference was generally perceived as making a valuable contribution to arts education in Alberta. Four issues, however, served to obfuscate the precise relationship of the conference purposes to the objectives of the Fine Arts Council. Consequently full assessment of the value of



the conference was hampered. These issues were:

- (i) vagueness of the objectives of the Fine Arts Council.
- (ii) unspecified relationship of conference purposes to the objectives of the Fine Arts Council.
- (iii) lack of a specific, overt, rationale for holding a joint "Fine Arts" conference.
- (iv) undefined "educative" role of the conference planners with respect to content selection.

#### Recommendations

The following recommendations are presented in the order they were derived from the evaluation questions.

- 1- The conference should be subject to evaluation again in 1976.

Questions were raised by participants with respect to a number of issues that were not clarified by this evaluation (due to the mail strike). The significant questions are presented on questionnaire form C (Appendix F).

Some major prior questions were raised by the evaluator concerning objectives of the Fine Arts Council, the function of the conference in relation to these objectives, the rationale for designation of the conference as a "Fine Arts" event and the "educative" role of the conference planning committee. These issues could well be more fully



explored in a further evaluation study.

2- The conference planning committee should examine  
publicity procedures.

(a) to determine if alternative procedures could be employed for wider distribution of publicity materials.

(b) to assess the feasibility of distributing detailed programs prior to the conference.

A number of registrants complained verbally about lack of publicity - 12 percent of respondents to the questionnaire recommended increased quantity of publicity.

A larger number of participants suggested prior distribution of the programs to assist with the decision to attend or not. Lessening conservatism with respect to session description may also serve to increase attendance.

3- Greater advantage could be made of the commercial  
displays.

(a) Display material appropriate for the drama group should be sought. No materials were available for this discipline in 1975.

(b) A suggestion was made by a participant, that a discussion session(s), based on the display materials, could be worthwhile. This may be a useful suggestion for future conferences.

4- Session content and scheduling should be re-examined.





(a) The possibility of repeating sessions could be examined. Twelve percent of respondents to the questionnaire made this suggestion. Flexibility of choice would consequently increase interest.

(b) Session content should be designated elementary, secondary or appropriate for both, to assist those participants with interest in specific levels. It should be remembered that teachers who attended were divided evenly between the elementary and secondary levels.

(c) The nature of the content should be re-examined in light of recommendation No. 5(d).

5- The four issues referred to in evaluation question No. 10 should be clarified by the executive and the conference planning committee.

These issues are stated below as prior questions that constitute important underpinnings of the conference rationale.

(a) What is meant by the objectives of the Fine Arts Council? Particularly troublesome phraseology includes:

- "To improve practice in the arts"
- "To improve members' knowledge and understanding (?) (of what?) in the fields of music, art and drama."
- "To promote an understanding of the functions of fine arts in the school and society". (It may be useful for the Council to state its position



with respect to the functions of fine arts in school and society.)

(b) What is the specific relationship of conference purposes to the objectives of the Fine Arts Council?

(c) To what extent is "interdisciplinary exchange" (as implied by the designation "Fine Arts Conference") a purpose of the conference?

(d) What is the responsibility of the conference planning committee with respect to content selection?

These questions were not observed to be considered overtly during the conference planning sessions, although certain shared assumptions were apparent. It is the opinion of the researcher that attention to the above questions should enable future conferences to make an even more powerful and purposeful contribution to arts education in Alberta.



## APPENDIX "D"



BROAD OBJECTIVES OF  
"INTRODUCING WAYS OF SEEING"

- OPENERS: Recognizing the existence of perceptual differences
- OBJECTIVE 1: Recognizing the visual as non-verbal communication
- Focus 1 Recognizing the communicative role of body language and spatial language
- Focus 2 Recognizing the communicative role of clothing and adornments
- Focus 3 Recognizing the communicative role of signs and symbols (discursive and non-discursive)
- OBJECTIVE 2: Investigating the elements of art
- Focus 4 Investigating line
- Focus 5 Investigating shape and space
- Focus 6 Investigating colour
- Focus 7 Investigating texture
- Focus 8 Investigating value
- OBJECTIVE 3: Investigating the principles of art
- Focus 9 Investigating balance
- Focus 10 Investigating dominance
- Focus 11 Investigating rhythm (repetition)
- Focus 12 Investigating variety
- Focus 13 Investigating unity





APPENDIX "E"



SELECTED SAMPLE OF LESSON OBJECTIVES AND  
ACTIVITIES FROM EPSB ART  
APPRECIATION PROJECT

AREA 1: Introducing Ways of Seeing

OBJECTIVE 2: Investigating the Elements of Art

FOCUS 6: Investigating Colour

To acquaint students with the concept of colour by exploring colour relationships such as: colour mixing, colour blending, comparing and contrasting colours.

---

ACTIVITY 1: With the help of your students organize a Colour Day. The purpose of such an event will be to single out one colour and enhance visual awareness of that colour as well to realize some of the moods, or emotions that colour may evoke over a prolonged period. Students could dress in that particular colour. A display could be arranged, using only that colour, taking in at least part of the environment. The art lesson in painting or drawing could be done in the chosen colour. In language arts, the children could build a word bank, i.e. - all the terms used to describe red could be charted. There



will of course be subject areas that do not tie in with the colour. In order not to depart from it completely, the children could do their written work in that colour or wear coloured goggles using cellophanes or gels. (The making of the goggles could be the first activity of the day - and might serve as a math lesson in measurement.)

### INTEGRATION:

Music and Language Arts - Tone colour in music and poetry can be compared. In music, voice and instruments can produce many identifiable sound qualities or "tone colours" e.g. strumming, plucking, scraping. In poetry, the tone colours imitate sounds, e.g. thud, crunch, crackle.

Refer to Silver Burdett's Music 3, pp.110-117  
Music 1, pp. 42- 45  
Music 4, pp.146-169  
 Relating The Arts, Gingrich, pp. 11-17

---

ACTIVITY 2: Have the students experiment with colour blends by giving them samples of colour (use coloured tissue, coloured gels, coloured cellophane, or coloured Playplax). Let them overlap their coloured samples and hold them



up to the light for blending. Older students can record their discoveries for future application.

NOTE:

This activity may be done on the overhead projector.

INTEGRATION:

Compare colour in painting with what is meant by colour in music and colour in word sounds. Refer to Relating The Arts, Donald Gingrich, pp. 9-17.

SCIENCE:

A lesson could be conducted in which students investigate why we see colour. (Experiment with colour absorption and reflection). See Film FI 0429- Understanding Colour.

ACTIVITY 4: Divide a 12" x 18" piece of paper into eight small sections by folding:



Now use crayons or paint to fill the sections in according to a scheme of opposites. These opposites could be based on; bright-dark,





cool-warm, delicate-strong, smooth-rough  
complementaries.

---

- ACTIVITY 5: Find variations of one colour from magazine pages. Encourage students to select a variety of hues and values in the chosen colour. The completed collage will be an example of a monochromatic colour scheme. Some guidelines for this particular exercise in collage:
- Have the students cut shapes rather than objects (a repetition of the same shape will contribute to unity)
  - A coloured background of the same chosen colour could be used
  - Try to match format with subject matter, i.e. reds in the shape of an apple, or spelling out the word red in large letters
  - The finished product will also indicate the ability of colour to produce texture.

TEACHER BACKGROUND NOTE:

COLLAGE is a composition made by gluing paper or other materials to an appropriate backing. (Derived from the french verb "coller", meaning to cut or paste).

MONOCHROMATIC SCHEME is one in which variations in tones of the same colour appear.



INTEGRATION:    Language Arts:

Have the students write their impressions of different colours: i.e. blue is ocean, sky, blueberries, etc. The format could be in prose, poetry or chart form.

---



APPENDIX "F"



## SUMMARY OF CONCLUSIONS

### EPSB "Art Appreciation" Evaluation

A statement of each evaluation question, discussed in Chapter VI preceeds the conclusions which relate to it.

1. How valuable were the outcomes of observed lessons in terms of the quality of experiences gained by the students?

All lessons were judged to provide some valuable experiences for students. Outcomes appeared to range from good to poor.

Valuable outcomes were as follows: i) comments and observations of students demonstrating responsiveness to visual experiences, ii) evidence of aesthetic judgements being made by students, iii) evidence of discoveries made by students that would contribute to visual awareness, iv) evidence of conditions which contributed to visually stimulating classroom environment, v) evidence of observations demonstrating visual awareness, vi) increasing willingness of students to discuss their work, vii) relatively high level of interest and enthusiasm.

Poor outcomes were observed to be: i) student action (making) without responding to visual materials, ii) discussions extended beyond the limit of student interests, iii) instructions issued to students which





limited their freedom of aesthetic choice in a significant way, iv) group discussion unrelated to the lesson activities.

Valuable outcomes were judged to outweigh the poor in terms of observed frequency and importance. Most of the undesirable outcomes were attributable to teaching style. One of the poor outcomes (i.e. students acting without responding to visual materials) was thought to be attributable to the nature of the exercise. (A list of similar exercises, together with an explanation of their limited usefulness was submitted to the curriculum developers).

Differences were noted between interpretations of the materials in Classrooms A and B across four dimensions. In specific terms these differences were reflected in

i) Linearity. Teachers in Classroom B specified more clearly the desired ends of lessons than did the teacher in Classroom A. This specification led to incidents of student frustration, although generally outcomes were viewed to be worthwhile.

ii) Orientation. Teachers in Classroom B, although professing attention to process, tended to stress product. This characteristic is related to linearity. Attention of students was consequently directed towards means of attaining alternate results rather than attending to the qualities existing in the product either during production or at its completion.

iii) Discussion. Student responses to work were



solicited more frequently in Classroom A than in Classroom B. Also the nature of discussions varied (as mentioned in ii).

iv) Continuity. Ideas were pursued from lesson to lesson in Classroom A so that a number of discussion "threads" were observed to develop. In Classroom B relationship between exercises was limited.

Each of these dimensions was judged to affect the value of outcomes. Although they are not directly attributable to the materials they could represent useful foci for evaluative examination following full implementation and in-service attention during implementation.

2. How well are non-art teachers able to direct art appreciation experiences given the open format of the guide materials?

Observation of one non-art classroom and discussion with twenty-one pilot teachers revealed widespread diffidence among non-art teachers with respect to directing discussions dealing with aesthetic qualities of visual objects. Lack of specificity of expected outcomes in the guide materials were viewed by these teachers to contribute to uncertainty with respect to discussions and judgements. Preference for a linear approach was also noted.

3. Do pilot teachers use the curriculum materials as intended by the developers?

Generally pilot teachers reported use of the



curriculum materials that was congruent with the developers' intentions. Some confusion was noted with respect to process versus product orientations.

4. How useful are the curriculum materials in the eyes of the pilot teachers.

The materials were regarded by all but two of the pilot teachers as a valuable source of ideas for art appreciation activities.

5. How does availability of visual aids affect the programme?

Most teachers observed that insufficient visual aids were available for most effective implementation of the programme. Although the materials listed EPSB Instructional Media Centre resources these were viewed to be so difficult to obtain at a predictable time that they were ineffective. Half the pilot teachers observed that lack of appropriate visual material could render the art appreciation programme ineffective.

6. How useful were the introductory meetings of the pilot group?

The introductory meetings of the pilot group were viewed by the majority of pilot teachers to be useful. Non-art teachers commented that these meetings were particularly significant in terms of assisting them with interpretation of the materials. Recommendations were made by pilot teachers for more practical sessions and "sharing" discussion meetings in the future.











**B30162**